Supplemental Appendix

A Appendix Tables

Table A.1: Change in Home Values over Distance from Central Business District and Distance from Highway (1950-1960, 1960-1970)

Variables	(1) OLS Δ Log Home Value	(2) IV for Log Dist Highway Plans	(3) Rays
Log Dist Highway	-0.0159	-0.0994***	-0.0469
	(0.0179)	(0.0373)	(0.0531)
Log Dist CBD	-0.168***	-0.148***	-0.161***
	(0.0389)	(0.0388)	(0.0410)
Redlined	0.284***	0.265***	0.277***
	(0.0648)	(0.0648)	(0.0667)
Dep. Var Mean (1960)	120,572 (2010$)		
R-squared	0.121	0.118	0.121
Observations	10,395	10,395	10,395
KP F-Stat		678.5	469.7

Notes: Observations are tracts from 1950, 1960, and 1970 (IPUMS NHGIS). The first difference is either over 1950 to 1960 or 1960 to 1970, depending on highway construction timing, and stacked into one panel. Limited to <5 miles of the nearest route and 30 miles from the central business district. Fixed effects at the CBSA level. Conley (1 km) standard errors reported. Control are for log distance from rivers, lakes, shores, ports, railroads, canals, and historical large roads, and the gradient (Dist CBD/Dist Highway). Redlined tracts are more than 80% HOLC D. Kleibergen-Paap rk Wald F statistics reported. *** p<0.01, ** p<0.05, * p<0.1

Table A.2: Residential Elasticity and Racial Preferences – Redlined Sample

Variables	(1) $\Delta \log L_{igr}$ (Δ Log Population 1960-1970) OLS	(2) + SES Cont
θ_r: $\Delta \log CMA_{igr}$		
Black	1.573***	1.482***
	(0.504)	(0.481)
	[0.633]	[0.623]
White	0.648***	0.627**
	(0.246)	(0.246)
	[0.396]	[0.399]
$\tilde{\rho}_r = \theta_r \rho_r$: $\Delta \log$ Pct White		
Black	-0.167*	-0.142
	(0.0984)	(0.0946)
	[0.131]	[0.131]
White	1.116***	1.108***
	(0.0719)	(0.0702)
	[0.0940]	[0.0928]
R-squared	0.322	0.337
Rounded Obs	56500	56500

Notes: Observations are first differences from 1960 to 1970 for tracts × group from the Census microdata. CBSA fixed effects by group included. Standard errors are cluster-robust by tract. Conley standard errors (1km) reported in brackets. The base controls are changes in log rent and 5 1-mile wide bins for distance from highways built between 1960-1970. SES controls are changes in log of income, pct high school graduate, pct top/bottom income quintile, and home values. Redlining fixed effects included in all specifications. The geographic controls are log distance from the CBD, rivers, lakes, shores, ports, historical railroads, canals, and historical large roads. All specifications include the Borusyak and Hull (2023) control for CMA in large roads. All controls are interacted with group. Observations are rounded to the nearest 500 for Census disclosure. *** p<0.01, ** p<0.05, * p<0.1

Table A.3: Observed vs. Predicted Outcomes Over Commuter Market Access Improvements

Variables	(1) Δ Log Obs Rent	(2) Δ Log Pred Rent	(3) Δ Log Obs Pct White	(4) Δ Log Pred Pct White	(5) Δ Log Obs Income	(6) Δ Log Pred Income
Δ Log CMA	0.0455***	0.0219***	0.00580	0.00639***	0.174***	0.172***
	(0.00626)	(0.000895)	(0.00392)	(0.000351)	(0.0219)	(0.00611)
R-squared	0.230	0.194	0.097	0.105	0.039	0.530
CBSA FE	Yes	Yes	Yes	Yes	Yes	
Geo Controls	Yes	Yes	Yes	Yes	Yes	
Rounded Obs	58000	58000	58000	58000	58000	58000

Notes: Unit of observation is census tract by race and education. Data comes from the first difference of 1960 to 1970 using restricted Census microdata. Fixed effects are at the CBSA (Core-based statistical area) level. Standard errors are cluster-robust with clusters at the tract-level. All specifications have as controls log distance from the central business district, rivers, lakes, shores, ports, historical railroads, canals, and historical large urban roads. Observation counts are rounded to the nearest 500 to meet Census disclosure rules. *** p<0.01, ** p<0.05, * p<0.1

Table A.4: Change in Log Productivity Over Distance from Highway (1960-1970)

Variables	(1) Δ Log Productivity
Dist from Highway (Miles)	-0.00269
	(0.00177)
Constant	-1.037***
	(0.0121)
R-squared	0.007
Rounded Obs	16000

Notes: Unit of observation is tract. Data comes from restricted Census microdata in 1960 and 1970. Standard errors are cluster-robust with clusters at the Place of Work Zone level because the variation in wages used to invert for productivity are at the Place of Work Zone level while housing prices used for inversion are at the tract level. Distance from the highway is in miles from segments of the highway network constructed between 1960 and 1970. Observation counts are rounded to the nearest 500 to meet Census disclosure rules. *** p<0.01, ** p<0.05, * p<0.1

Table A.5: Log Productivity and Pct HOLC D in 1960 and 1970

Variables	(1) Log Prod 1960	(2) Log Prod 1970
Pct HOLC D	0.0954	0.0358
	(0.0761)	(0.0719)
Dist from CBD (Miles)	0.00212	-0.00106
	(0.00171)	(0.00137)
Constant	5.797***	4.829***
	(0.0397)	(0.0332)
R-squared	0.020	0.011
Rounded Obs	17000	14000

Notes: Unit of observation is tract. Data comes from restricted Census microdata in 1960 and 1970. Standard errors are cluster-robust with clusters at the Place of Work Zone level because the variation in wages used to invert for productivity are at the Place of Work Zone level while housing prices used for inversion are at the tract level. Observation counts are rounded to the nearest 500 to meet Census disclosure rules. *** p<0.01, ** p<0.05, * p<0.1

Table A.6: Alternative Exercises for Welfare Changes (%) by Group

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Race	Black	Black	White	White	By Race		By Educ	
x Educ	<HS	HS+	<HS	HS+	Black	White	<HS	HS+
General Equilibrium								
Baseline	-1.45	-0.16	2.69	3.01	-1.04	2.86	2.07	2.79
Highway Impacts Separately								
Commuting Benefits	7.32	8.77	10.16	9.87	7.78	10.01	9.74	9.80
Localized Costs	-8.08	-8.05	-6.57	-6.03	-8.07	-6.28	-6.79	-6.17
Full Interstate Network								
Welfare Change	15.36	23.45	25.53	27.59	17.95	26.62	24.01	27.31
Alternative Road Placements								
Planned Routes	14.83	23.15	25.23	27.67	17.49	26.52	23.68	27.36
Euclidean Rays	21.32	30.14	32.31	34.74	24.14	33.60	30.67	34.43
Alternative Spillovers								
$\rho'_W = 0$	-1.44	-0.19	2.79	3.03	-1.04	2.92	2.16	2.81
$\rho'_B = -0.2$	-0.90	-0.01	2.70	3.03	-0.62	2.87	2.16	2.82
$\rho_r = 0, \gamma^A = 0$	-1.44	-0.21	2.79	3.00	-1.05	2.90	2.16	2.78
Alternative Elasticities								
$\theta'_B = \theta_W = 0.8$	-1.47	-0.19	2.52	2.86	-1.06	2.70	1.93	2.65
$\theta'_r = 3\theta_r$	-1.26	0.02	3.28	3.52	-0.85	3.41	2.60	3.28
$\theta'_B = \theta'_W = 3\theta_W$	1.02	1.01	2.67	2.94	1.02	2.81	2.42	2.81

Notes: Welfare calculations are based on data from the restricted Census in 1960. All welfare changes are for the general equilibrium simulation of highway impacts but with different parameter values. $\rho'_N = -0.2$ comes from the estimate for $\tilde{\rho}_N = -0.07$ in Table 3 divided by the residential elasticity of $\theta_N = 0.35$. The alternative elasticities set the residential elasticity for Black and White households to the same values at the level of White households, to three times their original values, to three times the original level of White households. All values are rounded to four significant digits to meet Census disclosure rules.

Table A.7: Changes in Equilibrium Outcomes (%) for Highway Impacts by Group

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Race	Black	Black	White	White	By Race		By Educ	
x Educ	<HS	HS+	<HS	HS+	Black	White	<HS	HS+
General Equilibrium								
Rent	0.32	0.47	0.23	0.14	0.37	0.18	0.24	0.16
Pct White	-0.47	-0.16	0.03	0.01	-0.37	0.02	-0.04	-0.00
Pct HOLC D	-0.19	-0.26	-1.06	-0.95	-0.21	-1.00	-0.93	-0.90
Amenities	-10.35	-9.43	-5.91	-4.82	-10.06	-5.33	-6.57	-5.13
Wages	-0.11	0.13	0.20	0.28	-0.03	0.24	0.15	0.27
Localized Costs	-7.99	-7.96	-6.37	-5.86	-7.98	-6.10	-6.61	-6.00
Dist from CBD	0.59	0.27	1.49	1.14	0.49	1.30	1.36	1.08
Commute Time	-3.72	-5.31	-3.47	-4.21	-4.23	-3.86	-3.51	-4.28
Commute Dist	4.81	3.95	6.16	4.91	4.53	5.50	5.96	4.84
General Equilibrium, No Barriers								
Rent	0.30	0.27	0.24	0.15	0.29	0.19	0.25	0.16
Pct White	0.09	0.08	-0.02	0.00	0.09	-0.01	-0.00	0.01
Pct HOLC D	-0.32	-0.29	-0.77	-0.77	-0.31	-0.77	-0.70	-0.74
Amenities	-6.00	-5.15	-6.17	-4.90	-5.73	-5.50	-6.14	-4.92
Wages	0.16	0.25	0.19	0.26	0.19	0.23	0.19	0.26
Localized Costs	-5.81	-5.80	-6.53	-5.99	-5.81	-6.24	-6.42	-5.98
Dist from CBD	0.67	0.47	1.52	1.18	0.61	1.34	1.39	1.13
Commute Time	-3.55	-4.16	-3.48	-4.21	-3.75	-3.87	-3.49	-4.21
Commute Dist	4.94	4.50	6.15	4.97	4.80	5.52	5.97	4.94

Notes: Equilibrium outcome calculations are based on data from the restricted Census in 1960. The general equilibrium simulation allows wages to respond in equilibrium. No institutions adjusts fundamental amenities for Black households by parameter E in redlined areas. The general equilibrium simulation with no institutions adds the highway impacts in the counterfactual world with no institutions. Parameter values are the same as in Table 4. All values are rounded to four significant digits for Census disclosure.

Table A.8: Changes in Welfare and Equilibrium Outcomes (%)
for Removal of Barriers by Group

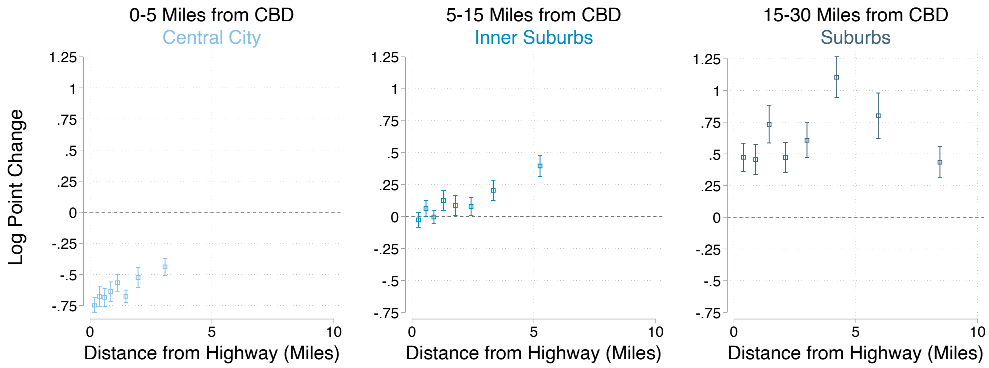
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Race	Black	Black	White	White	By Race		By Educ	
x Educ	<HS	HS+	<HS	HS+	Black	White	<HS	HS+
Welfare Change	—	—	-0.03	-3.27	—	-1.75	—	—
Rent	9.35	11.62	-0.56	-0.33	10.08	-0.44	0.92	0.48
Pct White	136.3	124.0	-5.04	-6.12	132.36	-5.61	16.02	2.75
Pct HOLC D	-52.74	-46.97	12.67	14.01	-50.89	13.38	2.93	9.85
Dist from CBD	93.37	94.16	-4.48	-3.25	93.62	-3.83	10.10	3.39
Amenities	—	—	-1.87	-5.64	—	-3.87	—	—

Notes: Equilibrium outcome calculations are based on data from the restricted Census in 1960. The simulation removes the location-race-specific wage for Black households, so welfare estimates are not available for them. Parameter values are the same as in Table 4. All values are rounded to four significant digits to meet Census disclosure rules.

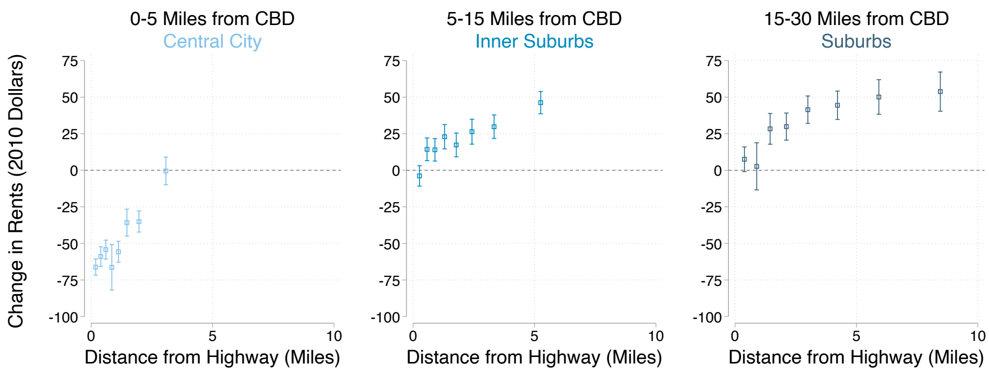
B Appendix Figures

Figure B.1: Changes Over Distance from Central Business District and Distance from Highway (1950-1960, 1960-1970)

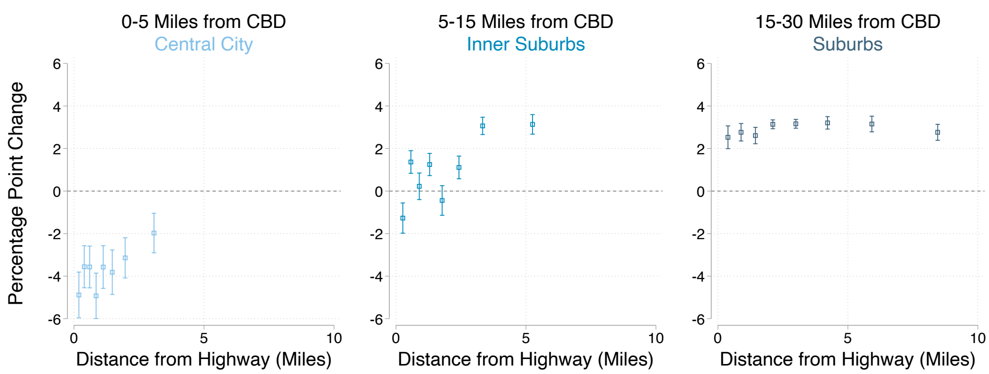
(a) Change in Log Population



(b) Change in Rents

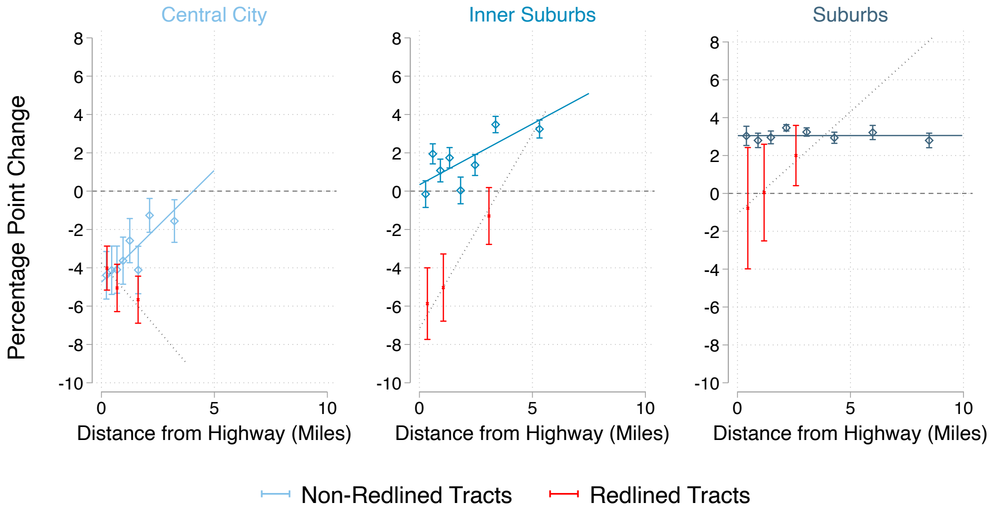


(c) Change in Percentage White



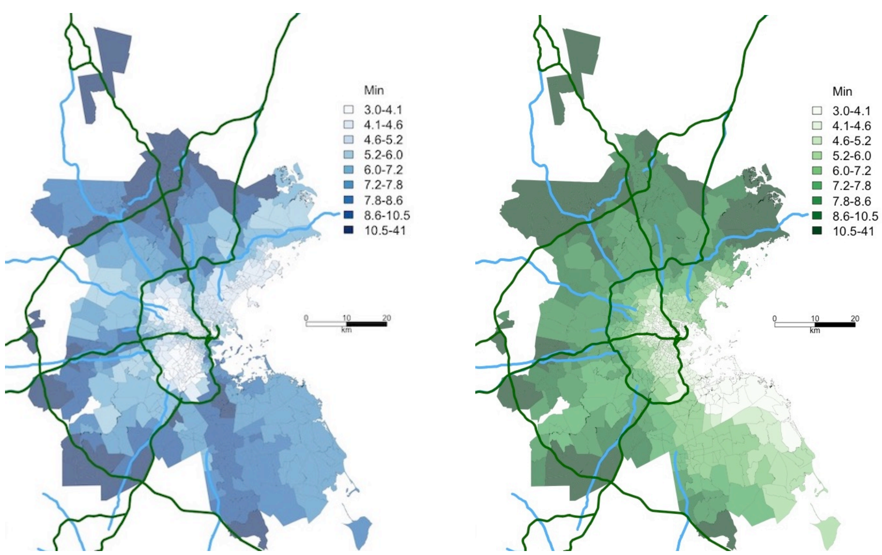
Notes: Unit of observation is census tract. Data comes from 1950, 1960, and 1970 tract-level aggregates retrieved from IPUMS NHGIS. The first difference is either over 1950 to 1960 or 1960 to 1970 and stacked into one panel depending on when highway construction started in the CBSA. All changes over time are de-meaned within CBSA. The sample of tracts for the central city panel is tracts within 5 miles of the constructed highway network, for the inner suburbs panel is tracts within 7.5 miles of the constructed highway network, and for the suburbs panel is tracts within 10 miles of the constructed highway network for legibility.

Figure B.2: Change in Percentage White by Redlining
Over Distance from Central Business District and Distance from Highway (1950-1960, 1960-1970)



Notes: Unit of observation is census tract. Data comes from 1950, 1960, and 1970 tract-level aggregates retrieved from IPUMS NHGIS. The first difference is either over 1950 to 1960 or 1960 to 1970 and stacked into one panel depending on when highway construction started in the CBSA. All changes over time are de-meaned within CBSA. The sample of tracts for the central city panel is those within 5 miles of the constructed highway network, for the inner suburbs panel is those within 7.5 miles of the constructed highway network, and for the suburbs panel is those within 10 miles of the constructed highway network for legibility. Redlined tracts are those where more than 80% of the area is redlined.

Figure B.3: Commute Time Reductions with Historical Roads as Interstates

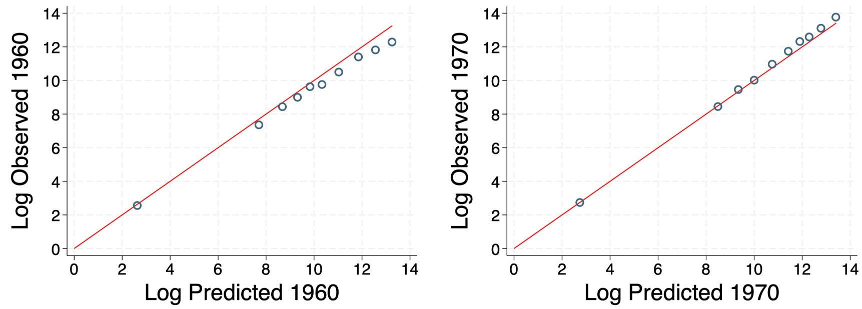


(a) Interstate Roads (b) Historical Large Roads as Interstates

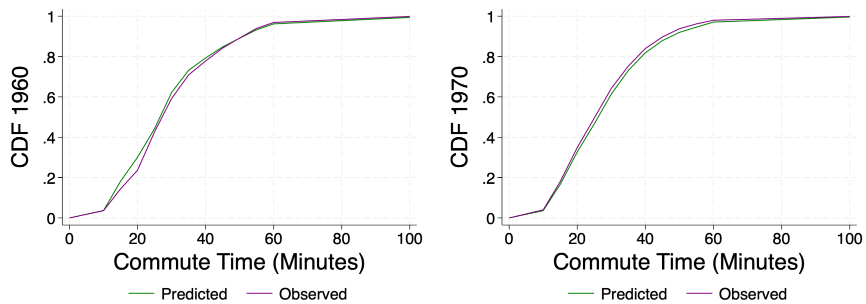
Notes: In Panel (a), commute time changes come from the author's calculations as the difference between commute times for the historical road network and for the entire Interstate network overlayed on the historical road network. In Panel (b) commute time changes come from the author's calculations as the difference between commute times for the historical road network and for the development of large roads as interstate highways.

Figure B.4: Predicted vs. Observed Commute Flows in 1960 and 1970

(a) Linear Fit and Scatter

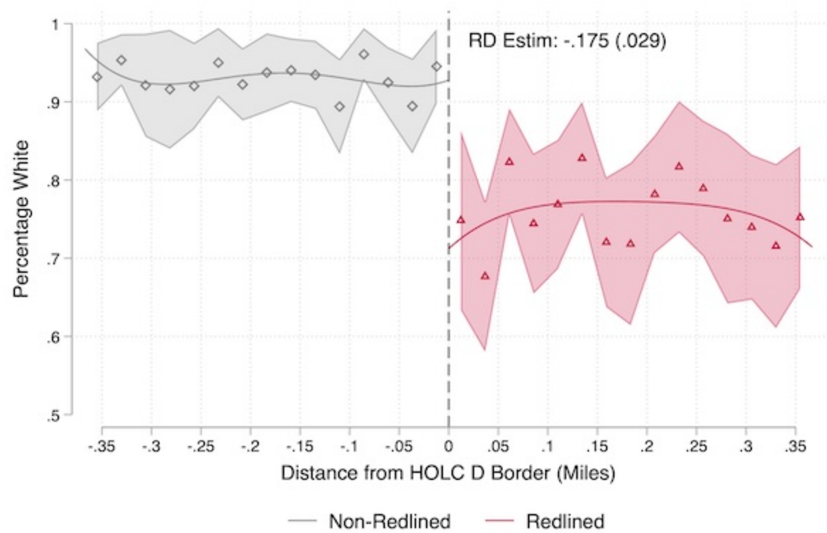


(b) Cumulative Distribution Function



Notes: Unit of observation is Place of Work Zone level by Place of Work Zone level. Data comes from restricted Census microdata in 1960 and 1970. In Panel (a), the scatter plot is created with 10 quantiles of predicted flows with analytical weights on the level of the observed commute flows. The red line is the 45 degree line. In Panel (b), the cumulative distribution function over commute time in minutes is in predicted flows for the green line and in observed flows for the purple line.

Figure B.5: Border Discontinuity for Percentage White in 1960 at HOLC D Border



Notes: Observations are 1960 census tract–HOLC border pairs (IPUMS NHGIS), comparing non-redlined (left) to redlined (right, ¿80% area is HOLC D). Each side has 15 bins and N=752. The RD uses an Epanechnikov kernel with optimal bandwidth 0.368 and a 4th-order polynomial Calonico et al. (2014). There are 15 bins on the left (N=752) and 15 bins on the right (N=752). The estimated coefficient is from the balanced sample RD shown in Appendix Table E.24 Panel B with order of polynomial=1, optimal bandwidth=0.368, and the same number of effective observations.

C Descriptive Results

C.1 Instrument Validity

To test for instrument validity, I examine pre-trends for changes between 1940 to 1950 or 1950 to 1960 depending on the timing of Interstate construction. As required for identification, the location of the planned routes and Euclidean rays is not correlated with demographic or economic changes before Interstate development, conditional on geographic controls. In Figure 3, there is also no cross-sectional correlation between the plans and rays with 1950 baseline characteristics after including controls.

Finally, I estimate the strength of the first-stage. To test that the planned routes and Euclidean rays are predictive of highway placement, I estimate two types of equations of the forms below where $IV \in \{Plans, Rays\}$.

$$\log(DistHW_i) = \phi \log(DistIV_i) + \mathbf{X}_i\theta + \lambda_{m(i)} + \nu_i$$
$$\mathbf{1}\{DistHW_i = 1\} = \pi\mathbf{1}\{DistIV_i = 1\} + \mathbf{X}_i\sigma + \delta_{m(i)} + \xi_i$$

The first compares log distance from the constructed routes to log distance from either the planned routes or the Euclidean rays, and the second compares a binary indicator for whether tracts are within 1 mile of the constructed route to the same indicator for the planned routes and rays to study placement at finer spatial scales. All the earlier controls and city fixed effects are included in the estimation. Results are shown in Table C.9. In both forms of first-stage regressions, the instruments are highly correlated with highway location as F-statistics on the excluded instrument are all above 100.

Table C.9: First-Stage for Highway Placement

Variables	(1) Log Dist HW	(2) Log Dist HW	(3) Dist HW = 1 mi	(4) Dist HW = 1 mi
Log Dist Plans	0.325***			
	(0.0175)			
Log Dist Rays		0.246***		
		(0.0196)		
Dist Plans = 1 mi			0.426***	
			(0.0223)	
Dist Rays = 1 mi				0.312***
				(0.0292)
Log Dist CBD	0.0863***	0.0752***	-0.0379***	-0.0422***
	(0.0212)	(0.0241)	(0.00957)	(0.0101)
F-Stat	342.8	156.5	366.3	113.9
R-squared	0.291	0.224	0.251	0.174
Observations	31,627	31,627	31,627	31,627
No. Counties	467	467	467	467

Notes: Unit of observation is census tract. Limited to those <5 miles of the nearest constructed route. CBSA fixed effects included. Standard errors are cluster-robust at the county level. All specifications control for log distance from the central business district, rivers, lakes, shores, ports, historical railroads, canals, and historical large roads. The reported F-stat comes from testing a single coefficient on the excluded instrument. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

C.2 Additional Results on Changes over CMA Improvements

(i) I estimate how the equilibrium outcomes of rents and racial composition respond to CMA in Appendix Table C.10 Columns 1 and 3. Consistent with the long differences, elasticities for rents and racial composition are smaller compared to the White population response and presumably play a smaller role in the welfare assessment. Because these equilibrium responses in turn affect population responses through feedback channels, I probe their importance for the population elasticities by controlling for rents and racial composition successively in Columns

2 and 4 as conducted in Adão et al. (2019). They do not appear to be a large determinant of the responses to CMA.

(ii) I construct two additional Borusyak and Hull (2023)-proposed CMA controls in Table C.11 Columns 1 and 2 where the planned routes and rays are converted into Interstates, and estimates remain unaffected when adding the controls. (iii) Pooling population elasticities by race is a fair approximation as in Table C.11 Column 4, I do not find substantial heterogeneity within race by education.

Table C.10: Elasticity of Rents, Pct White, and Population to Commuter Market Access

	(1)	(2)	(3)	(4)
Variables	Δ Log Rent	Δ Log Pop + Δ Log Rent Cont	Δ Log Pct White	Δ Log Pop + Δ Log Pct White Cont
$\Delta \log CMA_{igr}$	0.0432***		-0.0180	
	(0.00720)		(0.0139)	
Black		0.137		0.141
		(0.0974)		(0.0959)
White		1.267***		1.403***
		(0.118)		(0.114)
R-squared	0.225	0.121	0.071	0.146
CBSA FE	Yes	Yes	Yes	Yes
Geo Controls	Yes	Yes	Yes	Yes
Rounded Obs	59000	59000	60000	60000

Notes: Unit of observation is census tract by race and education. Data comes from the first difference of 1960 to 1970 using restricted Census microdata. Fixed effects are at the CBSA (Core-based statistical area) level. Standard errors are cluster-robust with clusters at the tract-level. The geographic controls are log distance from the central business district, rivers, lakes, shores, ports, historical railroads, canals, and historical large urban roads, all interacted with race in Columns 2 and 4 and with race and education in Columns 1 and 3. Observation counts are rounded to the nearest 500 to meet Census disclosure rules. All specifications include the Borusyak and Hull (2023) control for CMA in large roads. Kleibergen-Paap rk Wald and Cragg-Donald Wald F statistics for weak instruments are reported. *** p<0.01, ** p<0.05, * p<0.1

C.3 Instrument for CMA

In Appendix Table C.12, I report the first-stage regressions between CMA and corresponding measures with the Interstate, plan, and ray instruments. Because CMA includes both wage and commute cost changes, the first-stage coefficients on the planned and ray CMA instruments are lower than the first-stage coefficients reported for placement in Table C.9. However, when wages are fixed to 1960 levels, the first-stage coefficients for CMA look similar to those for placement as then the variation only comes from Interstate highway construction.

D Quantitative Model

D.1 Model Extensions

D.1.1 Stone-Geary

An alternative approach to incorporating non-homotheticity in housing consumption is to allow for Stone-Geary preferences. The consumer maximization problem then includes a minimum amount of housing consumption $\bar{l}_{gr}$

Table C.11: Elasticity of Population to Commuter Market Access – Additional Results

	(1)	(2)	(3)
	$\Delta \log L_{igr}$ (Δ Log Population 1960-1970)		
Variables	+BH (2023) Plans	+BH (2023) Rays	Unscaled CMA CMA
$\Delta \log CMA_{igr}$			
Black	0.104	0.107	
	(0.0973)	(0.0970)	
White	1.469***	1.458***	
	(0.121)	(0.121)	
$\Delta \log \Phi_{igr}$			
Black			0.0314
			(0.0323)
White			0.470***
			(0.0385)
R-squared	0.113	0.113	0.113
CBSA FE	Yes	Yes	Yes
Geo Controls	Yes	Yes	Yes
Rounded Obs	N = 60500		

Notes: Unit of observation is census tract by race and education. Data comes from the first difference of 1960 to 1970 using restricted Census microdata. Fixed effects are at the CBSA (Core-based statistical area) level. Standard errors are cluster-robust with clusters at the tract-level. The geographic controls are log distance from the central business district, rivers, lakes, shores, ports, historical railroads, canals, and historical large urban roads, all interacted with race. Column 1 and Column 2 include the Borusyak and Hull (2023) control for CMA interacted with race when the planned network and the Euclidean ray network are built, respectively. Column 3 includes the Borusyak and Hull (2023) control for CMA in large roads interacted with race. Observation counts are rounded to the nearest 500 to meet Census disclosure rules. *** p<0.01, ** p<0.05, * p<0.1

Table C.12: First-Stage for Commuter Market Access Improvements

	(1)	(2)	(3)	(4)	(5)
Variables	Δ Log CMA	Δ Log CMA	Δ Log CMA	Δ Log CMA HW	Δ Log CMA HW
Δ Log CMA HW	0.639***				
	(0.0127)				
Δ Log CMA Plans		0.107***		0.382***	
		(0.0120)		(0.0088)	
Δ Log CMA Rays			0.0888***		0.336***
			(0.0108)		(0.0079)
F-Stat	2534	79.68	67.14	1884	1797
R-squared	0.313	0.262	0.262	0.505	0.484
CBSA FE	Yes	Yes	Yes	Yes	Yes
Geo Controls	Yes	Yes	Yes	Yes	Yes
Rounded Obs	60500	60500	60500	60500	60500

Notes: Unit of observation is census tract by race and education. Data comes from the first difference of 1960 to 1970 using restricted Census microdata. Fixed effects are at the CBSA (Core-based statistical area) level. Standard errors are cluster-robust with clusters at the tract-level. All specifications have as controls log distance from the central business district, rivers, lakes, shores, ports, historical railroads, canals, and historical large urban roads. All specifications include the Borusyak and Hull (2023) control for CMA in large roads. Observation counts are rounded to the nearest 500 to meet Census disclosure rules. The reported F-stat comes from testing a single coefficient on the excluded instrument. *** p<0.01, ** p<0.05, * p<0.1

that can differ by group.

$$\max_{c_{ij}(o),l_i(o)} \quad \frac{z_i(o)\varepsilon_j(o)B_{igr}}{d_{ijgr}}\left(\frac{c_{ij}(o)}{\beta_{gr}}\right)^{\beta_{gr}}\left(\frac{l_i(o)-\bar{l}_{gr}}{1-\beta_{gr}}\right)^{1-\beta_{gr}}$$
$$\text{s.t.} \quad c_{ij}(o)+Q_i l_i(o)=\frac{w_{jgr}\varepsilon_j(o)}{d_{ijgr}}$$

which leads to the indirect utility function of

$$u_{ijgr}(o)=B_{igr}z_i(o)Q_i^{\beta_{gr}-1}\left(\frac{w_{jgr}\varepsilon_j(o)}{d_{ijgr}}-Q_i\bar{l}_{gr}\right)$$

Stone-Geary allows for income changes to generate sorting in contrast to Cobb-Douglas preferences.

D.1.2 Nested Frechet

With the indirect utility function defined as before, suppose that $z_i(o)$ is instead distributed with a nested Frechet structure where the cumulative distribution is

$$F(z_i(o))=\exp\Big(-\Big(\sum_n\Big(\sum_{i\in S_n} z_i(o)^{-\theta_r}\Big)^{-\lambda_r/\theta_r}\Big)\Big)$$

where λ_r governs the substitutability across types of neighborhoods. When $\lambda_r=\theta_r$, the expression returns to the usual Frechet distribution as before. In this setting, suppose $\lambda_B<\theta_B$ where for Black households, there is more substitutability within type of neighborhood than across, and $\lambda_W=\theta_W$ where for White households, their choice behavior is not nested across the type of neighborhoods.

The choice probabilities are then

$$\begin{aligned}\pi_{igr}&=\pi_{n,gr}\pi_{igr|n}\\&=\frac{\left(\sum_{s\in S_n}\left(B_{sgr}CMA_{sgr}Q_s^{\beta_{gr}-1}\right)^{\theta_r}\right)^{\lambda_r/\theta_r}}{\sum_m\left(\sum_{s\in S_m}\left(B_{sgr}CMA_{sgr}Q_s^{\beta_{gr}-1}\right)^{\theta_r}\right)^{\lambda_r/\theta_r}}\frac{\left(B_{igr}CMA_{igr}Q_i^{\beta_{gr}-1}\right)^{\theta_r}}{\sum_{s\in S_n}\left(B_{sgr}CMA_{sgr}Q_s^{\beta_{gr}-1}\right)^{\theta_r}}\end{aligned}$$

following a two step process. First, there is a choice of type n of neighborhoods, and then conditional on type, there is a choice of neighborhood i within type n.

Define $V_{n,gr}=\left(\sum_{s\in S_m}\left(B_{sgr}CMA_{sgr}Q_s^{\beta_{gr}-1}\right)^{\theta_r}\right)^{1/\theta_r}$ as the inclusive value of living in type n neighborhoods. The share living in type n follows the usual gravity share formula with the shape parameter λ_r.

$$\pi_{n,gr}=\frac{V_{n,gr}^{\lambda_r}}{\sum_m V_{m,gr}^{\lambda_r}}$$

The population elasticity to CMA $\frac{\partial \pi_{igr}}{\partial CMA_{igr}}$ using the product rule and the definition of $\pi_{igr} = \pi_{n,gr}\pi_{igr|n}$ is

$$\frac{\partial \pi_{igr}}{\partial CMA_{igr}} = \frac{\partial \pi_{n,gr}}{\partial CMA_{igr}}\pi_{igr|n} + \frac{\partial \pi_{igr|n}}{\partial CMA_{igr}}\pi_{n,gr}$$
$$\frac{\partial \pi_{igr}}{\partial CMA_{igr}} = \frac{\lambda_r \pi^2_{igr|n}}{CMA_{igr}}\pi_{n,gr}(1-\pi_{n,gr}) + \frac{\theta_r \pi_{n,gr}}{CMA_{igr}}\pi_{igr|n}(1-\pi_{igr|n})$$
$$\Rightarrow \frac{\partial \pi_{igr}}{\partial CMA_{igr}}\frac{CMA_{igr}}{\pi_{igr}} = \lambda_r \pi_{igr|n}(1-\pi_{n,gr}) + \theta_r(1-\pi_{igr|n})$$

When $\lambda_r = \theta_r$ (as is the case for White households), then the elasticity becomes

$$\begin{aligned}\frac{\partial \pi_{igr}}{\partial CMA_{igr}}\frac{CMA_{igr}}{\pi_{igr}} &= \theta_r \pi_{igr|n}(1-\pi_{n,gr}) + \theta_r(1-\pi_{igr|n}) \\ &= \theta_r(\pi_{igr|n} - \pi_{igr}) + \theta_r(1-\pi_{igr|n}) = \theta_r(1-\pi_{igr})\end{aligned}$$

which is the same as when there are no nests for types.

When $\lambda_r < \theta$ (as is the case for Black households), then the elasticity is lower

$$\lambda_r \pi_{igr|n}(1-\pi_{n,gr}) + \theta_r(1-\pi_{igr|n}) < \theta_r(1-\pi_{igr})$$

even though the conditional elasticity (conditional on type of neighborhood) is still approximately θ.

$$\frac{\partial \pi_{igr|n}}{\partial CMA_{igr}}\frac{CMA_{igr}}{\pi_{igr|n}} = \theta_r(1-\pi_{igr|n})$$

D.2 Separate Idiosyncratic Shocks

Two separate idiosyncratic shocks are received for residences and workplaces. Residential shocks $z_i(o)$ are drawn from distribution $F(z_i(o)) = \exp(-z_i(o)^{-\theta_r})$ and workplace shocks $\varepsilon_j(o)$ are likewise distributed Frechet from $F(\varepsilon_j(o)) = \exp(-T_{jgr}\varepsilon_j(o)^{-\phi})$.

Previous estimates of the combined shock leverage variation on the workplace side, so in this model, ϕ represents the substitution elasticity across workplaces. The model implies residential choice follows the equation $L_{igr} = (B_{igr}\Phi_{igr}^{\frac{1}{\phi}}Q_i^{\beta_{gr}-1})^{\theta_r} / \sum_t (B_{tgr}\Phi_{tgr}^{\frac{1}{\phi}}Q_t^{\beta_{gr}-1})^{\theta_r}\mathbb{L}_{gr}$ where $\Phi_{igr} = \sum_j T_{jgr}(w_{jgr}/d_{ijgr})^{\phi}$. Under the null hypothesis that the residential elasticity and workplace elasticity are equivalent $\phi = \theta_r$, the coefficient λ_r in the estimating equation $\log L_{igr} = \lambda_r \log \Phi_{igr} + \gamma_{gr}$ should be approximately 1. Note that in Φ_{igr}, no assumptions are made on the value of ϕ because in the commuting gravity equation, $\nu_{gr} = \kappa_{gr}\phi$ is estimated directly from the data. See Appendix E.1 for details on the values in Φ_{igr}. In Appendix Table C.11, I test for whether the elasticities to residential and workplace shocks should be the same value. I find the coefficient on Φ_{igr} is significantly less than one, suggesting the residential elasticity is in fact lower than the labor supply elasticity to workplaces.

D.3 Spatial Barriers and Isomorphisms

In the indirect utility function, an amenity wedge is isomorphic to a price wedge according to the relationship

$$1 - \tau^b_{igr} = \left(1 + \tau^Q_{igr}\right)^{\beta_{gr}-1}$$

so to attain the capacity constraint, the pride wedge is then $\tau^Q_{igr} = \mathbf{k}_{igr}^{1/(\theta_r(\beta_{gr}-1))} - 1$.

To achieve a capacity constraint $\bar{c}_{igr}$, as the constraint becomes tighter i.e. $\bar{c}_{igr} \to 0$, the barriers to entry for a neighborhood become larger. An equivalent way to arrive at the same allocation is to sufficiently increase the amenity or price wedge. The amenity wedge, when no price wedge exists, must satisfy

$$\frac{L_{igr}}{\mathbb{L}_{gr}} = \frac{\bar{c}_{igr}}{\mathbb{L}_{gr}} = \frac{\left((1-\tau^b_{igr})B_{igr}CMA_{igr}Q_i^{\beta_{gr}-1}\right)^{\theta_r}}{\sum_{t\neq i}\left(B_{igr}CMA_{igr}Q_i^{\beta_{gr}-1}\right)^{\theta_r} + \left((1-\tau^b_{igr})B_{igr}CMA_{igr}Q_i^{\beta_{gr}-1}\right)^{\theta_r}}$$

$$\Rightarrow \frac{\bar{c}_{igr}}{\mathbb{L}_{gr}} \sum_{t\neq i}\left(B_{igr}CMA_{igr}Q_i^{\beta_{gr}-1}\right)^{\theta_r} = \left(1-\frac{\bar{c}_{igr}}{\mathbb{L}_{gr}}\right)\left((1-\tau^b_{igr})B_{igr}CMA_{igr}Q_i^{\beta_{gr}-1}\right)^{\theta_r}$$

$$\Rightarrow \frac{\bar{c}_{igr}/\mathbb{L}_{gr}\sum_{t\neq i}\left(B_{tgr}CMA_{tgr}Q_t^{\beta_{gr}-1}\right)^{\theta_r}}{(1-\bar{c}_{igr}/\mathbb{L}_{gr})\left(B_{igr}CMA_{igr}Q_i^{\beta_{gr}-1}\right)^{\theta_r}} = (1-\tau^b_{igr})^{\theta_r} = \mathbf{k}_{igr} \Rightarrow \tau^b_{igr} = 1-\mathbf{k}_{igr}^{1/\theta_r}$$

The average of the idiosyncratic shocks of individuals in each location can be derived using the properties of the Frechet distribution and generally follows $\bar{z}_{igr} = \Gamma\left(1-\frac{1}{\theta_r}\right)\pi_{igr}^{1/\theta_r}$. Substituting $\mathbf{k}_{igr}$ and residential factors into π_{igr} leads to the expression for average shocks.

$$\bar{z}_{igr} = \Gamma\left(1-\frac{1}{\theta_r}\right)\left(\frac{\sum_{t\neq i}\left(B_{tgr}CMA_{tgr}Q_t^{\beta_{gr}-1}\right)^{\theta_r} + \mathbf{k}_{igr}\left(B_{tgr}CMA_{tgr}Q_t^{\beta_{gr}-1}\right)^{\theta_r}}{\mathbf{k}_{igr}\left(B_{tgr}CMA_{tgr}Q_t^{\beta_{gr}-1}\right)^{\theta_r}}\right)^{1/\theta_r}$$

Note that while a capacity constraint, amenity wedge, and price wedge lead to the same allocation, the general equilibrium implications do differ. Suppose capacity constraints only bind for Black households such that $L_{igB} = \bar{c}_{igB}$ while the White population is determined endogenously and responds to neighborhood-level changes. If the capacity constraint is implemented via a price wedge, the housing market is subsequently affected by changes in housing demand across groups, versus via an amenity wedge, it is not directly impacted.

Concretely, the price wedge lowers housing demand (consumption). This then changes the expression for total housing supply, which is shared across all groups and is equated with total housing consumption.

$$H_i = \sum_g H_{igB} + \sum_g H_{igW} = \sum_g \frac{(1-\beta_{gB})\bar{w}_{igB}L_{igB}}{(1+\tau^Q_{igB})Q_i} + \sum_g \frac{(1-\beta_{gW})\bar{w}_{igW}L_{igW}}{Q_i} \tag{12}$$

D.4 Firms and Housing

Firms – As workers alter their labor supply to workplaces in response to reductions in commute costs, wages are determined in equilibrium by firms. While adjustments at firms are not a central theme of the empirical evidence or the question of the paper, I include this feature to close the model and allow for a comprehensive assessment of the impacts of Interstate highways. In the counterfactual exercises, I probe its importance for welfare by shutting down firm adjustments in wages and housing.

Across workplaces, there are representative firms with constant returns to scale production so that demand by firms translates into demand at each workplace. Perfectly competitive firms produce varieties with Cobb-Douglas technology over labor and commercial floorspace following $Y_j = A_j N_j^{\alpha} H_{Fj}^{1-\alpha}$, where α is the share of labor and A_j is a Hicks-neutral productivity shock. Combining heterogeneous workers, labor N_j is a CES aggregate over education where workers of different education levels are imperfect substitutes (Katz and Murphy, 1992; Card, 2009). N_{jg} is further a CES aggregate of different racial groups.

$$N_j = \left(\sum_g \alpha_{jg} N_{jg}^{\frac{\sigma^g - 1}{\sigma^g}} \right)^{\frac{\sigma^g}{\sigma^g - 1}} \quad \text{with} \quad N_{jg} = \left(\sum_r \alpha_{jgr} L_{Fjgr}^{\frac{\sigma^r - 1}{\sigma^r}} \right)^{\frac{\sigma^r}{\sigma^r - 1}}$$

This nested-CES structure accommodates imperfect substitutability across race as Boustan (2009) finds Black workers are closer substitutes to each other than to White workers. Imperfect substitutability can arise from occupational segregation preventing workers from switching into an occupation predominantly of another race or from unobserved skill gaps, even conditional on education (Higgs, 1977).[36]

Within education, locations employ workers from each race at varying intensities α_{jgr}, incorporating how firms in the central city may have different demands for Black workers compared to firms in the suburbs e.g. due to discrimination across space (Holzer and Reaser, 2000; Miller, 2023). Moreover, it generalizes the labor aggregate structure of Tsivanidis (2023) by exploiting the detailed workplace wage Census data.

Firm profit maximization generates labor and commercial floorspace demand with the corresponding wage indices.

$$L_{Fjgr} = \left(\frac{w_{jgr}}{\alpha_{jgr}\omega_{jg}} \right)^{-\sigma^r} \left(\frac{\omega_{jg}}{\alpha_{jg} W_j} \right)^{-\sigma^g} N_j \quad \text{s.t.} \;\; W_j = \left(\sum_g \alpha_{jg}^{\sigma^g} \omega_{jg}^{1-\sigma^g} \right)^{\frac{1}{1-\sigma^g}} \quad \omega_{jg} = \left(\sum_r \alpha_{jgr}^{\sigma^r} w_{jgr}^{1-\sigma^r} \right)^{\frac{1}{1-\sigma^r}} \tag{13}$$

$$H_{Fj} = \left(\frac{1-\alpha}{Q_j} A_j \right)^{1/\alpha} N_j \tag{14}$$

The zero-profit condition from perfect competition combined with profit maximization leads to the subsequent condition for commercial rental prices, which rise when productivity is high and wages are low. Firms thus aim to locate in more productive, cheaper, and lower-wage areas.

$$Q_j = (1-\alpha) \left(\frac{\alpha}{W_j} \right)^{\frac{\alpha}{1-\alpha}} A_j^{\frac{1}{1-\alpha}} \tag{15}$$

Agglomeration – Within productivity of locations, the term A_j contains a fundamental component a_j that does not vary with equilibrium outcomes and an endogenous component representing agglomeration economies in density (L_{Fj}/K_j). In the definition of density, L_{Fj} is total employment, K_j is the land area, and γ^A is the strength of agglomeration.

$$A_j = a_j \left(L_{Fj}/K_j \right)^{\gamma^A} \tag{16}$$

Transportation infrastructure in conjunction with agglomeration can reallocate economic activity as in Faber (2014), Heblich et al. (2020), and Baum-Snow (2020) with racially disparate effects as studied by Miller (2023). The model includes this channel of density affecting productivity (Rosenthal and Strange, 2004; Ellison et al., 2010).

Housing – Given that empirically, housing prices adjusted with highway construction, I allow for a housing construction sector that responds elastically to changes in demand from both residences and workplaces. In each location, there is H_i amount of floorspace that is allocated endogenously across residential versus commercial uses where θ_i is the share for residential use. Residential floorspace demand aggregates across the housing expenditures

[36] The average Black worker at this time attended lower quality schools, especially in the segregated South, compared to the average White worker which would lead equivalent years of schooling to translate into different skill levels (Margo, 2007).

of each group Exp_{igr} so H_{Ri} is determined following

$$H_{Ri} = \theta_i H_i = \sum_{g,r} \frac{Exp_{igr}}{Q_i} \text{ with } Exp_{igr} = (1-\beta_{gr})\bar{w}_{igr}\varphi_{gr}L_{igr},\ \bar{w}_{igr} = \sum_j \pi_{j|igr}w_{jgr} \tag{17}$$

Distribution of rents to homeowners is constant across neighborhoods for each group within each city. Let total income by group gr be the sum of total labor income and total rental income from housing rents to each group based on the share of home values that the group owns in the portfolio of the city.

$$\underbrace{\varphi_{gr}\sum_i \bar{w}_{igr}L_{igr}}_{\text{total income}} = \underbrace{\sum_i \bar{w}_{igr}L_{igr}}_{\text{total labor income}} + \underbrace{\sum_i \hat{o}_{igr}\sum_{g,r} Exp_{igr}}_{\text{total rental income}}$$
$$\Rightarrow \varphi_{gr} = 1 + \frac{\sum_i \hat{o}_{igr}\sum_{g,r} Exp_{igr}}{\sum_i \bar{w}_{igr}L_{igr}}$$

The share of home values $\hat{o}_{igr}$ is observed in the data as the proportion of home values in homes owned by group gr out of total home values in a neighborhood.

Commercial floorspace demand comes from firm optimization in Eq. (14), and with the two expressions for residential and commercial floorspace demand, the allocation across uses θ_i and total floorspace demand $H_i = H_{Ri} + H_{Fi}$ are then determined for land market clearing.

To parameterize how housing is supplied elasticity, I follow the literature where the housing production function is $H_i = K_i^{\mu} M_i^{1-\mu}$ with M_i as capital at universal price p and K_i as land at price r_i (Epple et al., 2010; Combes et al., 2021). The implied supply curve is

$$H_i = \left(\frac{1-\mu}{\mu}\right)^{\frac{1-\mu}{\mu}} K_i Q_i^{\frac{1-\mu}{\mu}} \tag{18}$$

D.5 General Equilibrium

Given the model's parameters $\{\beta_{gr}, \theta_r, \kappa_{gr}, \phi, \alpha, \alpha_{jg}, \alpha_{jgr}, \sigma^g, \sigma^r, \mu, \rho_r, \gamma^A\}$, city populations by education and race $\{\mathbb{L}_{gr}\}$, and location characteristics $\{T_{jgr}, t_{ijgr}, b_{igr}, a_j, K_i\}$, the general equilibrium is represented by the vector of endogenous objects $\{L_{igr}, L_{Fjgr}, Q_i, \theta_i, w_{jgr},$
$B_{igr}, A_j, U_{gr}\}$ determined by the following equations:

1. *Residential populations* in each neighborhood (5)
2. *Labor supply* at each workplace (3)
3. *Housing demand* from residences and firms (17) + (14)
4. *Housing supply* from the construction sector (18)
5. *Zero profit and profit maximization* by firms (13)
6. *Endogenous amenities* from racial composition (6)
7. *Endogenous productivity* from agglomeration (16)
8. *Closed City* where $\sum_i L_{igr} = \mathbb{L}_{gr}$

E Inversion and Estimation

E.1 Model Inversion

Parameters Estimated During Model Inversion

- $\alpha_{jg}, \alpha_{jgr}$ are labor intensities for the CES nested labor aggregate

- T_{jgr} is the scale parameter for workplaces

Observed data sources

- Observed wages $\hat{w}_{jgr}$ come from the Decennial microdata

Step 1 – Given $\{L_{Fjgr}, L_{igr}, t_{ijgr}\}$ and the semi-elasticity of commuting parameter $\{\nu_{gr}\}$, I invert for composite transformed wages $\omega_{jgr} = T_{jgr} w_{jgr}^{\phi}$ from the labor supply equation following

$$\begin{aligned} L_{Fjgr} &= \sum_i \frac{T_{jgr}(w_{jgr}/d_{ijgr})^{\phi}}{\sum_s T_{sgr}(w_{sgr}/d_{is})^{\phi}} L_{igr} \\ &= \sum_i \frac{\omega_{jgr}/\exp(\nu_{gr} t_{ijgr})}{\sum_s \omega_{sgr}/\exp(\nu_{gr} t_{isgr})} L_{igr} \end{aligned}$$

Commuting costs are in terms of commute times t_{ijgr} following $d_{ijgr} = t_{ijgr}^{\kappa_{gr}}$, therefore $d_{ijgr}^{\phi} = t_{ijgr}^{\nu_{gr}}$ with $\nu_{gr} = \kappa_{gr}\phi$. Labor supply is in the second line rewritten as a function of composite transformed wages ω_{jgr}. Wages are solved for iteratively following the process of Ahlfeldt et al. (2015) and wages are only identified up to a scaling factor.

Step 2 – Given $\{\omega_{jgr}\}$, the Frechet shape parameter for labor supply ϕ, and observed wages $\{\hat{w}_{jgr}\}$, I back out the Frechet scale parameter T_{jgr}. Following that $\omega_{jgr} = T_{jgr} w_{jgr}^{\phi}$, then the Frechet scale parameter is $T_{jgr} = \omega_{jgr}/w_{jgr}^{\phi}$. Compared to existing work where wages are not directly observed, this additional data allows for separately identifying the workplace amenity value T_{jgr} from the scale wages component ω_{jgr}.

Residential Side

Step 3 – Given $\{Q_i, \omega_{jgr}, t_{ijgr}, L_{igr}\}$ and the parameters $\{\beta_{gr}, \phi, \kappa_{gr}, \theta_r\}$, I can recover residential amenities B_{igr}. Returning to the residential choice equation, the share of each demographic group that lives in a location i follows

$$\frac{L_{igr}}{\mathbb{L}_{gr}} = \frac{\left(B_{igr} CMA_{igr} Q_i^{\beta_{gr}-1}\right)^{\theta_r}}{\sum_t \left(B_{tgr} CMA_{tgr} Q_{tr}^{\beta_{gr}-1}\right)^{\theta_r}}$$

which can be rearranged using the welfare equation (7).

$$\left(\frac{L_{igr}}{\mathbb{L}_{gr}}\right)^{1/\theta_r} = \frac{B_{igr} CMA_{igr} Q_i^{\beta_{gr}-1}}{U_{gr}}$$

Choosing units for amenities such that the geometric mean $\bar{B}_{igr} = \left[\prod_{i=1}^{S} B_{igr}\right]^{1/S} = 1$, and continuing with the bar notation for geometric mean, I calibrate amenities following

$$\frac{B_{igr}}{\bar{B}_{igr}} = \left(\frac{L_{igr}}{\bar{L}_{Rigr}}\right)^{1/\theta_r} \left(\frac{Q_i}{\bar{Q}_{ir}}\right)^{1-\beta_{gr}} \left(\frac{CMA_{igr}}{\overline{CMA}_{igr}}\right)^{-1}$$

Workplace Side

Step 4 – Given $\{L_{Fjgr}, w_{jgr}\}$ and the parameters $\{\sigma^r, \sigma^g\}$, I estimate the parameters α_{jg}, α_{jgr} with the following procedure. Using the labor demand equation from (13), the share of labor employed in a location and in an education group g that is of race r is

$$\frac{L_{Fjgr}}{L_{Fjg}} = \frac{(w_{jgr}/\alpha_{jgr})^{-\sigma^r}}{\sum_s (w_{jgs}/\alpha_{jgs})^{-\sigma^r}}$$

The share of labor and wages are observed, so this equation allows for determining α_{jgr} with the constraint that $\sum_r \alpha_{jgr} = 1$.

With a similar process, I solve for α_{jg}. First, I calculate $N_{jg} = \left(\sum_r \alpha_{jgr} L_{Fjgr}^{\frac{\sigma^r-1}{\sigma^r}}\right)^{\frac{\sigma^r}{\sigma^r-1}}$ which is a function of observed and previously estimated values. Using the CES demand form, I arrive at the equation

$$\frac{N_{jg}}{\sum_h N_{jh}} = \frac{(\omega_{jg}/\alpha_{jg})^{-\sigma^g}}{\sum_h (\omega_{jh}/\alpha_{jh})^{-\sigma^g}}$$

which is an equation for the unknown α_{jg} with the constraint that $\sum_g \alpha_{jg} = 1$. Recall that $\omega_{jg} = \left(\sum_r \alpha_{jgr}^{\sigma^r} w_{jgr}^{1-\sigma^r}\right)^{\frac{1}{1-\sigma^r}}$, which is a function of known values.

Step 5 – Given $\{q_i, w_{jgr}\}$ and the parameters $\{\alpha, \alpha_{jg}, \alpha_{jgr}\}$, I recover workplace productivity A_i. Productivity for each location i is inferred from the zero profit equation.

$$q_i = (1-\alpha)\left(\frac{\alpha}{W_j}\right)^{\frac{\alpha}{1-\alpha}} A_i^{\frac{1}{1-\alpha}} \text{ for } i \in Tracts_j$$

where $W_j = \left(\sum_g \alpha_{jg}^{\sigma^g} \omega_{jg}^{1-\sigma^g}\right)^{\frac{1}{1-\sigma^g}}$ the price index for labor is calculated after backing out wages w_{jgr} and the $\alpha_{jg}, \alpha_{jgr}$ relative productivity parameters at the POW zone j. Since prices are observable at the tract level for tract i, I assume that wages are the same for all tracts in POW zone j which is the set $Tracts_j$.

Housing Supply and Allocation

Step 6 – Given $\{Q_i, \omega_{jgr}, t_{ijgr}, L_{igr}, q_i, A_j, L_{Fjgr}\}$, the parameters $\{\beta_{gr}, \phi, \kappa_{gr}, \alpha, \alpha_{jg}, \alpha_{jgr}\}$, I recover total housing supply H_i and floorspace allocation θ_i across commercial and residential uses. Returning to the residential and commercial demand for floorspace equations, we have for the residential side

$$H_{Ri} = \theta_i H_i = \sum_{g,r} \frac{Exp_{igr}}{Q_i} \text{ with } Exp_{igr} = (1-\beta_{gr})\bar{w}_{igr}\varphi_{gr}L_{igr},\ \bar{w}_{igr} = \sum_j \pi_{j|igr} w_{jgr}$$

Distribution of rents to homeowners is calculated using total income by group gr as the sum of total labor income and total rental income to each group based on the share of home values that the group owns in the portfolio of the city.

$$\underbrace{\varphi_{gr}\sum_i \bar{w}_{igr}L_{igr}}_{\text{total income}} = \underbrace{\sum_i \bar{w}_{igr}L_{igr}}_{\text{total labor income}} + \underbrace{\sum_i \hat{o}_{igr}\sum_{g,r} Exp_{igr}}_{\text{total rental income}}$$
$$\Rightarrow \varphi_{gr} = 1 + \frac{\sum_i \hat{o}_{igr}\sum_{g,r} Exp_{igr}}{\sum_i \bar{w}_{igr}L_{igr}}$$

For the commercial side, I use that the production function is Cobb-Douglas

$$H_{Fi} = \left(\frac{W_j}{\alpha}\right)\left(\frac{1-\alpha}{q_i}\right)\frac{N_j}{S_j} \text{ for } i \in Tracts_j$$
$$\text{where } N_j = \left(\sum_g \alpha_{jg} N_{jg}^{\frac{\sigma^g-1}{\sigma^g}}\right)^{\frac{\sigma^g}{\sigma^g-1}} \text{ and } N_{jg} = \left(\sum_r \alpha_{jgr} L_{Fjgr}^{\frac{\sigma^r-1}{\sigma^r}}\right)^{\frac{\sigma^r}{\sigma^r-1}}$$

As the geographic unit on the workplace side is a POW zone while the geographic unit on the residential side is a census tract, I assume that on the workplace side, labor is supplied evenly across all the tracts within a POW zone.

Finally, θ_i where i is at the tract-level is set to follow

$$\theta_i = \frac{H_{Ri}}{H_i} = \frac{H_{Ri}}{H_{Ri} + H_{Fi}}$$

with housing supply at the tract level $H_i = H_{Ri} + H_{Fi}$.

Step 7 – Given $\{H_i, Q_i\}$ and the parameter μ, I recover the scaled amount of land used for development k_i as a location fundamental following profit maximization of the construction sector. Demand for capital can be derived as $M_i = Q_i H_i (1-\mu)/p$. Substituting this equation into the housing production function gives

$$H_i = K_i Q_i^{\frac{1-\mu}{\mu}} \left(\frac{1-\mu}{p}\right)^{1-\mu}$$
$$H_i = k_i Q_i^{\frac{1-\mu}{\mu}} \qquad \text{where } k_i = K_i \left(\frac{1-\mu}{p}\right)^{1-\mu}$$

In the quantitative implementation, I allow μ (the capital intensity of housing construction) which determines the housing supply elasticity to price to differ in the suburbs versus the central city following recent work by Baum-Snow and Han (2021) and Saiz (2010). Let $\mu = 0.3$ for neighborhoods < 5 miles of the CBD and $\mu = 0.2$ for neighborhoods ≥ 5 miles from the CBD, corresponding to floorspace supply elasticities of 2.33 and 4, respectively.

E.2 Parameter Estimation

E.2.1 Estimation of discriminatory pricing

Discriminatory pricing will be measured directly from home values and rents in the microdata. To test for differential pricing by race across neighborhoods, I look at the coefficient from the interaction of race and redlining. The estimating equation is across observations for each household h with either log home value or log rent as the dependent variable.

$$\log Q_h = \alpha_i + \alpha_r + \phi_1 D_i^{red} + \phi_2 D_i^{red} \times D_h^{non-white} + X_h + \varepsilon_h$$

α_i is for neighborhood fixed effects, α_r is for race fixed effects, D_i^{red} is a dummy for being in a redlined neighborhood, D_h^{black} is a dummy for the household head being Black, and X_h is a set of household level characteristics on the quality of the home such as the availability of air conditioning, a freezer, a toilet, or a bathtub. The coefficient ϕ_2 is the differential increase in price black households have to pay to live outside of redlined neighborhoods compared to white households. In Table E.13, it appears that Black households pay less than White households for similar quality housing in non-redlined neighborhoods and more in redlined neighborhoods. These results do not suggest pricing is the reason why Black households are more likely to live in redlined areas.

Table E.13: Housing Price Discrimination in Rents and Home Values in 1960

Variables	*Panel A* – Log Rent (1)	(2) Tract FE	(3) Quality	(4) Tract+Qual	*Panel B* – Log Home Value (1)	(2) Tract FE	(3) Quality	(4) Tract+Qual
Black	-0.155***	-0.0250***	-0.0295***	0.0260***	-0.332***	-0.0453***	-0.143***	-0.0146**
	(0.0117)	(0.00719)	(0.00939)	(0.00584)	(0.0175)	(0.0112)	(0.0121)	(0.00675)
Redlined	-0.340***		-0.212***		-0.361***		-0.205***	
	(0.0107)		(0.00804)		(0.0171)		(0.0134)	
Black × Redlined	0.183***	0.0845***	0.0929***	0.0508***	0.142***	0.0602***	0.0967***	0.0509***
	(0.0174)	(0.0103)	(0.0138)	(0.00873)	(0.0319)	(0.0189)	(0.0245)	(0.0123)
Constant	4.272***				9.625***			
	(0.00514)				(0.00567)			
R-squared	0.104	0.433	0.394	0.592	0.078	0.522	0.404	0.658
Rounded Obs	1,729,000	1,729,000	1,729,000	1,729,000	1,562,000	1,562,000	1,562,000	1,562,000

Notes: Unit of observation is household. Household level data comes from the 1960 Census microdata. Fixed effects are at the census tract level. Standard errors are cluster-robust with clusters at the tract level. The quality controls include categorical variables for availability of air conditioning, dryer, elevator, freezer, hot water, kitchen, shower, basement, toilet, and the type of heating, type of fuel for cooking, type of fuel for heat, type of fuel for water, source of water, source of water, sewage facilities, number of stories, number of rooms, number of bathrooms, number of bedrooms, and year built. Redlined tracts are tracts where more than 80% of the area is redlined. Observation counts are rounded to nearest 1000 to meet Census disclosure rules. *** p<0.01, ** p<0.05, * p<0.1

E.2.2 Gravity Equation

Using the commute shares in Eq. (2) and the functional form of commute costs as $d_{ijgr} = t_{ijgr}^{\kappa_{gr}}$, I estimate the following gravity equation for commuting elasticities.

$$\log \pi_{j|igr,t} = \underbrace{\gamma_{jgr,t}}_{\log \omega_{jgr}} + \underbrace{\gamma_{igr,t}}_{\log \Phi_{igr}} - \underbrace{\nu_{gr}}_{\kappa_{gr}\phi} \log t_{ijgr,t} + \varepsilon_{ijgr,t}$$

Location by year fixed effects $\gamma_{jgr,t}$ and $\gamma_{igr,t}$ account for factors that are workplace-specific (scaled wages ω_{jgr}) and residence-specific (transformed commuter access Φ_{igr}) in each year. The error term $\varepsilon_{ijgr,t}$ captures remaining factors outside of the model or mismeasurement in commute times. Commute flows from the 1960 and 1970 Censuses are pooled together, and commute times are computer-generated. Bilateral variation in commute times then identifies the elasticity. These times are instrumented using the plans and rays for the post-highway period.

Splitting by race and education leads to some zero-count bilateral pairs, which happens often for the Black population (11 percent of the sample). To reduce sparsity, I aggregate residential tracts up to the Place of Work Zone, so estimation is for POW Zone by POW Zone by year with standard errors clustered for POW Zone by POW Zone. Even with the aggregation, some bilateral pairs continue to have zero counts for the Black population. In addition to estimating the log-log specification above, I conduct the robustness checks suggested by the trade literature in Head and Mayer (2014) and estimate the commuting elasticity with Poisson Pseudo Maximum Likelihood (PPML) following Silva and Tenreyro (2006) to address sparsity.

In Table E.14, I find that less-educated groups, both White and Black, tend to have higher elasticities compared to the higher-educated. Parameter values for Black workers are lower than values for White workers, suggesting that Black households consider commute times less in their commuting decisions. These values are similar to those in Heblich et al. (2020) of -4.90, estimated with commuting data from 19th-century London. First stages for the instruments are reported in Supp. Appendix Table E.15, and the estimates from Panel B are the preferred values used for the quantitative analysis. For the Black population, the commuting elasticity rises slightly with PPML estimates in Panel C. For White workers, their elasticities are lowered with the PPML estimator although the observation count only increases a small amount.[37] Overall, the pattern remains quite similar. Lastly, in Panels D and E, I instrument the PPML estimates via a control function approach following Wooldridge (2015) and bootstrap standard errors. These values concur with the previous PPML estimates, so instrumenting is not crucial.

Table E.14: Commuting Gravity Equation

Race x Educ	(1) Black <HS	(2) Black HS Grad	(3) White <HS	(4) White HS Grad
$\nu_{gr} = \kappa_{gr}\phi$	*Panel A* – Log Commuting Share – IV Plans			
Log Commute Time	-4.206***	-3.671***	-4.707***	-4.168***
	(0.126)	(0.120)	(0.0673)	(0.0505)
R-squared	0.232	0.182	0.377	0.367
	Panel B – Log Commuting Share – IV Rays			
Log Commute Time	-4.197***	-3.645***	-4.708***	-4.154***
	(0.127)	(0.122)	(0.0674)	(0.0503)
R-squared	0.232	0.182	0.377	0.367
Rounded Obs	7000	8000	21500	25000
	Panel C – Poisson Pseudo Maximum Likelihood			
Log Commute Time	-4.703***	-3.929***	-3.877***	-3.247***
	(0.0819)	(0.0599)	(0.0471)	(0.0359)
	Panel D – Commuting Share (PPML) – IV Plans			
Log Commute Time	-4.706***	-3.940***	-3.888***	-3.260***
	(0.138)	(0.0857)	(0.0655)	(0.0526)
	Panel E – Commuting Share (PPML) – IV Rays			
Log Commute Time	-4.707***	-3.941***	-3.883***	-3.256***
	(0.140)	(0.0879)	(0.0655)	(0.0522)
Rounded Obs	20500	21000	26000	27000

Notes: Unit of observation is Place of Work Zone by Place of Work Zone pair by year where commuting flows from residential tracts are aggregated up to the Place of Work Zone geography. Data comes from the restricted Census microdata in 1960 and 1970. Fixed effects are for POR (Place of Residence) by year at the Place of Work Zone unit (although it represent residence, not workplace). POW by year fixed effects are also included for workplaces at the POW Zone level. The conditional commuting share is the share from a residential location that commutes to a workplace. The observation counts are lower for the Black population as some residences and workplaces have zero Black population (while PPML addresses zeros in bilateral flows, it does not address zeros in entire rows or columns). Standard errors are cluster-robust with clusters at the Place of Work Zone by Place of Work Zone level. For the IV-PPML estimates, to obtain standard errors, I bootstrap 200 samples with clusters at the Place of Work Zone by Place of Work Zone level. Observation counts are rounded to 500 for Census disclosure. *** p<0.01, ** p<0.05, * p<0.1

[37] Accounting for zeros in bilateral pairs still leaves the observation count of the less-educated Black population at 22000 below that of the higher-educated White population as there are cases of residential and workplace units without any Black workers, and PPML only adjusts for *bilateral* counts of zero.

Table E.15: Commuting Gravity Equation – Additional Results

	(1)	(2)	(3)	(4)
Race x Educ	Black <HS	Black HS Grad	White <HS	White HS Grad
	Panel A – First-Stage – IV Plans			
Log Commute Time	0.988***	1.026***	1.023***	1.025***
	(0.00573)	(0.00475)	(0.00178)	(0.00145)
F-Stat (Rounded)	29710	46750	331400	501900
	Panel B – First-Stage – IV Rays			
Log Commute Time	0.999***	1.037***	1.027***	1.029***
	(0.00595)	(0.00492)	(0.00183)	(0.00152)
F-Stat (Rounded)	28150	44440	315400	455600

Notes: Observations are POW Zone by POW Zone pairs by year where commuting flows from residential tracts are aggregated up to the POW Zone. Data is the restricted Census microdata in 1960 and 1970. Fixed effects are for POW Zones by year for residences and separately for workplaces. The conditional commuting share is the share of each residence commuting to a workplace. Counts are lower for the Black population as some residences and workplaces have zero Black population. Standard errors are cluster-robust by POW Zone $\times$ POW Zone. Observation rounded to the nearest 500 for Census disclosure. The F-stat tests a single coefficient on the excluded instrument and is rounded to four significant digits for Census disclosure. *** p<0.01, ** p<0.05, * p<0.1

E.2.3 Linear prediction of housing consumption share from CEX data

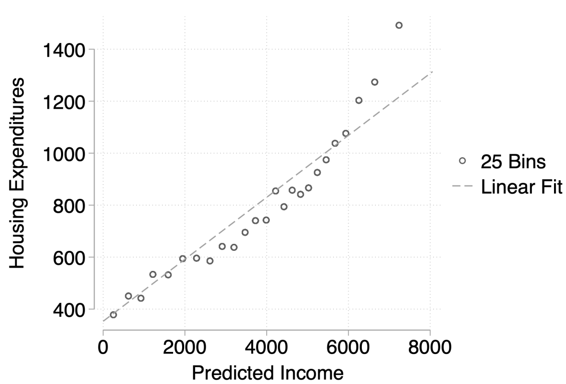
I assign the housing consumption share for each race by education by first estimating a linear function for housing expenditure over income from the Consumer Expenditure Surveys Public-Use Microdata in the year 1980.

$$E_i^{hous} = \beta_{gr}^0 + \beta_{gr}^1 PredIncome_i + \varepsilon_i$$

E_i^{hous} is quarterly housing expenditure and $PredIncome_i$ is quarterly income predicted using categorical variables in age, education, marital status, occupation, sex, race and region. I use predicted rather than observed income given the variability in observed income that would lead to downward biased estimates of β_{gr}^1.

As depicted in Figure E.6, the assumption of linearity for the housing consumption Engel curve seems plausible. From this function, I calculate the predicted housing expenditure for each group, given their average income, and set the ratio of predicted housing expenditure to average income as the housing consumption share.

Figure E.6: Linear Housing Expenditure Function Over Income



Notes: Unit of observation is individual. Data comes from the Consumer Expenditure Surveys Public-Use Microdata in 1980. Predicted income is a linear prediction of income using categorical variables in age, education, marital status, occupation, sex, race and region. Income and housing expenditure is for quarterly amounts. 25 equally sized bins in predicted income (when predicted income is greater than zero) are created for the scatter. The linear fit uses the estimated coefficients from Table E.16.

Table E.16: Housing Expenditure Function

	(1)
Variables	Housing Expenditures
Predicted Income	0.119***
	(0.00294)
Constant	353.3***
	(9.263)
R-squared	0.080
Observations	20,786

Notes: Unit of observation is individual. Data comes from the Consumer Expenditure Surveys Public-Use Microdata in the year 1980. Predicted income is a linear prediction of income using categorical variables in age, education, marital status, occupation, sex, race and region. Income and housing expenditure is for quarterly amounts. Standard errors are heteroskedasticity robust. *** p<0.01, ** p<0.05, * p<0.1

Table E.17: Predicted Share of Income Spent on Housing

Race x Educ	Black <HS	Black HS Grad	White <HS	White HS Grad
Housing Exp. Share	0.34	0.22	0.30	0.21
Observations	1441	1908	5885	14712

Notes: Unit of observation is individual. Data comes from the Consumer Expenditure Surveys Public-Use Microdata in the year 1980. Income and housing expenditure is in quarterly amounts. Predicted share of income spent on housing uses the linear housing expenditure function from E.16 and the average level of income of the four race by education groups.

E.2.4 Localized Costs

After estimating the residential elasticity and racial preference parameters, I invert the model to recover fundamental amenities b_{igr} and estimate how the Interstate highway system affected nearby neighborhoods' amenities following $b_{igr} = 1 - b^{HW}\exp(-\eta DistHW_i)$.

$$\Delta \log b_{igr} = \sum_{k=1}^{5} \beta_k \mathbf{1}\{DistHW_i = k\} + \mathbf{X_i}\eta_{gr} + \gamma_{m(i)gr} + \alpha_{red(i)} + \varepsilon_{igr}$$

The exponential decay is approximated with non-parametric mile-wide bins up to 5 miles from Interstate segments built between 1960 and 1970. The equation controls for distance from the CBD and geographic features in $\mathbf{X_i}$ interacted with group, city by group fixed effects, and redlining fixed effects. Standard errors are clustered at the tract level.

In Column 1 of Table E.18 with no controls, there is a large drop in fundamental amenities where at 1 mile from the constructed network, $\Delta \log B_{igr} = -0.453$ (0.0501). However, much of the decline by highways is due to selection in route placement as including geographic controls in Column 2 reduces the estimate to -0.119 (0.0516). To gain precision, I further report results for 0.5 mile-wide bins in Column 3 where $\Delta \log b_{igr} = -0.191$

(0.0581) in the first 0.5 mile. These estimates are comparable in size to findings in Brinkman and Lin (2022) using cross-sectional variation from Chicago. To assign parameter values for b^{HW} and η, I match the functional form of $b_{igr} = 1 - b^{HW}\exp(-\eta DistHW_i)$ to two of the estimated bins in Column 3 at $k = 0.5, 1.5$.[38]

The population response is due to not just the direct negative consequences of highways but also the indirect changes in racial composition. I therefore set the composite amenity term $B_{igr} = b_{igr}(L_{iW}/L_i)^{\rho_r}$ as another outcome in Column 4 to include the indirect amenity impacts. I find endogenous amenities explain a small portion of the population drop by highways since the estimated value is only slightly more negative at $\Delta \log B_{igr} = -0.124$ (0.0517). Instrumented results are shown in Table E.19 and are too noisy to measure values precisely. In Table Appendix E.20, I project modern-day measures of environmental pollution over distance from Interstate roads and find a 2% increase within the first mile, which is a strong lower bound on pollution during Interstate construction as from the 1960s to today, car pollutants have been reduced by 99%.[39]

In a falsification test, I measure whether there were fundamental amenity changes near other historical large roads. Roads may have universally become more congested or polluted, and declines near Interstate highways may not be a distinctive feature that should be counted fully for welfare impacts. I replace the distance bins from Interstate roads with distance bins from historical control roads that were never re-built $\{\mathbf{1}\{DistLARGE_i = k\}_{k=1,\ldots,5}\}$. Falsification results in Column 4 indicate no change in amenities near large roads with the estimate at 1 mile being very close to zero at -0.0057 (0.137). The negative consequences are thus a unique aspect of Interstate routes where their massive size and elevated ramps were particularly unpleasant for neighboring areas (Rose and Mohl, 2012).

E.2.5 External Parameters

In the production function, the labor share is set to 0.7 following findings in Greenwood et al. (1997). The elasticity of substitution by education σ^g in the CES labor aggregate comes from Card (2009) which uses the education categories of high school versus college educated. Estimates range from 1.4 to 3 and are corroborated by several other sources, so I set $\sigma^g = 2$ (Borjas, 2003; Ottaviano and Peri, 2012). The elasticity of substitution by race is taken from Boustan (2009) to be $\sigma^r = 8$. Housing supply elasticity values are obtained from Baum-Snow and Han (2021) and set to differ in the central city (within 5 miles of the CBD) where $\mu_{cbd} = 0.35$ versus the suburbs (all other neighborhoods) where $\mu_{sub} = 0.25$. Lastly, the agglomeration parameter is set to 0.07 within the range of Rosenthal and Strange (2004) and Kline and Moretti (2014).

Table E.21: Additional Model Parameters

Parameters	Source
Production Labor Share	
$\alpha = 0.7$	Greenwood et al. (1997)
Elasticity of Substitution by Race and Education	
$\sigma^r = 8$, $\sigma^g = 2$	Card (2009), Boustan (2009)
Agglomeration	Rosenthal and Strange (2004),
$\gamma^A = 0.07$	Kline and Moretti (2014)
Housing Supply Elasticity	
$\mu^{cbd} = 0.35$, $\mu^{sub} = 0.25$	Baum-Snow and Han (2021)

Notes: Values are set following the literature.

[38] Parameters are set to solve two equations: $1 - b^{HW}\exp(-\eta 0.5) = \exp(-0.191), 1 - b^{HW}\exp(-\eta 1.5) = \exp(-0.0994)$.

[39] The Clean Air Act of 1970 was the first of many federal legislative efforts to reduce air pollution.

Table E.18: Change in Amenities over Distance from Highway

	(1)	(2)	(3)	(4)	(5)
		$\Delta \log b_{igr}$		$\Delta \log B_{igr}$	$\Delta \log b_{igr}$
Variables	OLS	OLS	0.5 mi bins	OLS	Placebo
Dist Highway (mi = 1)	-0.453***	-0.119**		-0.124**	0.00565
	(0.0501)	(0.0516)		(0.0517)	(0.137)
Dist Highway (mi = 2)	-0.379***	-0.0933*		-0.125**	-0.0707
	(0.0499)	(0.0515)		(0.0515)	(0.0997)
Dist Highway (mi = 3)	-0.223***	0.000345		-0.0343	-0.0752
	(0.0531)	(0.0552)		(0.0553)	(0.0910)
Dist Highway (mi = 4)	-0.0795	0.0824		0.0458	0.0307
	(0.0596)	(0.0604)		(0.0606)	(0.0899)
Dist Highway (mi = 5)	0.0369	0.143**		0.140**	0.0437
	(0.0642)	(0.0638)		0.0638)	(0.0793)
Dist Highway (mi = 0.5)			-0.191***		
			(0.0581)		
Dist Highway (mi = 1)			-0.0651		
			(0.0572)		
Dist Highway (mi = 1.5)			-0.0994*		
			(0.0588)		
Dist Highway (mi = 2)			-0.0888		
			(0.0573)		
Dist Highway (mi = 2.5)			0.0116		
			(0.0618)		
Dist Highway (mi = 3)			-0.0153		
			(0.0667)		
Dist Highway (mi = 3.5)			0.0703		
			(0.0738)		
Dist Highway (mi = 4)			0.0955		
			(0.0731)		
Dist Highway (mi = 4.5)			0.140*		
			(0.0808)		
Dist Highway (mi = 5)			0.146*		
			(0.0807)		
R-squared	0.028	0.052	0.052	0.050	0.069
CBSA X Group FE	Yes	Yes	Yes	Yes	Yes
Geo Controls		Yes	Yes	Yes	Yes
Rounded Obs	49500	49500	49500	49500	9000

Notes: Observations are the first difference of 1960 to 1970 for census tracts by race and education from the restricted Census microdata. CBSA by race and education fixed effects included. Standard errors are cluster-robust by tract. There are 5 binary indicators for distance from highways built between 1960 and 1970 in 1-mile wide bins. In Column 3, the bins are split further into 0.5-mile wide bins. The geographic controls are log distance from the CBD, rivers, lakes, shores, ports, historical railroads, canals, and historical large roads, all interacted with race and education. Limited to tracts $<$ 10 miles of highway routes. For the placebo test in Column 5, the sample is restricted to $>$ 5 miles of a highway and $<$ 10 miles of historical large roads. The control for distance from historical large roads is dropped since it is now the endogenous variable. All specifications include redlining fixed effects. Observation counts are rounded to the nearest 500 for Census disclosure rules. Parameters used to invert for dependent variables are in Table 5. *** p$<$0.01, ** p$<$0.05, * p$<$0.1

Table E.19: Change in Amenities over Distance from Highway – IV

	(1)	(2)
	$\Delta \log b_{igr}$	
	IV – 2 mi bins	
Variables	Dist Plans	Dist Rays
Dist Highway (mi = 2)	-0.0469	0.519
	(0.133)	(0.668)
Dist Highway (mi = 4)	0.249*	0.312
	(0.149)	(0.502)
Dist Highway (mi = 6)	0.0766	-0.149
	(0.279)	(1.385)
R-squared	0.047	0.036
CBSA X Group FE	Yes	Yes
Geo Controls	Yes	Yes
Rounded Obs	51000	47000
C-D F-Stat	686.2	18.38
K-P F-stat	110.8	4.55

Notes: Observations are the first difference of 1960 to 1970 for census tracts by race and education from the restricted Census microdata. CBSA by race and education fixed effects included. Standard errors are cluster-robust by tract. There are 3 binary indicators for distance from highways built between 1960 and 1970 in 2-mile wide bins to increase power. The geographic controls are log distance from the central business district, rivers, lakes, shores, ports, historical railroads, canals, and historical large urban roads, all interacted with race and education. Redlining fixed effects are included. Limited to tracts $<$ 10 miles of planned routes or the Euclidean ray network. Observation counts are rounded to 500 for Census disclosure. Kleibergen-Paap rk Wald and Cragg-Donald Wald F statistics reported. Parameters used to invert for dependent variables are in Table 5. *** p<0.01, ** p<0.05, * p<0.1

Table E.20: Environmental Pollution Index (PM 2.5) over Distance from Highway

	(1)	(2)	(3)
	Log Particulate Matter 2.5		
Variables	1 mi bins	+ Geo Cont	0.5 mi bins
Dist Highway (mi = 1)	0.0245***	0.0204***	
	(0.000983)	(0.000976)	
Dist Highway (mi = 2)	0.0231***	0.0196***	
	(0.000980)	(0.000965)	
Dist Highway (mi = 3)	0.0197***	0.0174***	
	(0.00102)	(0.00100)	
Dist Highway (mi = 4)	0.0146***	0.0133***	
	(0.00114)	(0.00111)	
Dist Highway (mi = 5)	0.0108***	0.0103***	
	(0.00129)	(0.00126)	
Dist Highway (mi = 0.5)			0.0201***
			(0.00107)
Dist Highway (mi = 1)			0.0207***
			(0.00101)
Dist Highway (mi = 1.5)			0.0200***
			(0.00103)
Dist Highway (mi = 2)			0.0191***
			(0.00106)
Dist Highway (mi = 2.5)			0.0176***
			(0.00110)
Dist Highway (mi = 3)			0.0172***
			(0.00117)
Dist Highway (mi = 3.5)			0.0145***
			(0.00130)
Dist Highway (mi = 4)			0.0119***
			(0.00136)
Dist Highway (mi = 4.5)			0.0118***
			(0.00154)
Dist Highway (mi = 5)			0.00850***
			(0.00167)
Dep Var Mean	13.50		
R-squared	0.962	0.964	0.964
CBSA FE	Yes	Yes	Yes
Geo Controls		Yes	Yes
Observations	32,833	32,833	32,833

Notes: Unit of observation is census tract. Fixed effects are at the CBSA (Core-based statistical area) level. Data comes from the CDC Environmental Health Census Tract-Level PM2.5 Concentrations, 2001-2005 measures. There are 5 binary indicators for distance from highways in 1-mile wide bins in Columns 1–2 and 10 binary indicators in 0.5-mile wide bins in Column 3 (the value displayed is the upper end of the bin). Included in all specifications are redlining fixed effects and log distance from the central business district. The geographic controls are log distance from rivers, lakes, shores, ports, historical railroads, canals, and historical large urban roads. The sample is limited to tracts within 10 miles of highway routes. Standard errors are heteroskedasticity robust. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

E.2.6 Instruments for Estimation of Endogenous Amenities/Preferences

There are two endogenous variables $\{\Delta \log CMA_{igr}, \Delta \log(L_{iW}/L_i)\}$ that require instruments. The instrument for $\Delta \log CMA_{igr}$ uses commute times where the plans or ray network are converted into Interstate highways and scaled wages are fixed to 1960 levels. Specifically, the CMA IV is defined as $CMA^{IV}_{igr} = \frac{1}{\phi}\left(\log \sum_j \omega_{jgr,1960}/d^{IV}_{ijgr}{}^{\phi}\right)$ $-\frac{1}{\phi}\left(\log \sum_j \omega_{jgr,1960}/d_{ijgr,1960}{}^{\phi}\right)$ with commute times in the post period from the planned or ray network. The the instruments for racial composition changes $\Delta \log(L_{iW}/L_i)$ are described below.

Davis Instruments – Instruments following Davis et al. (2019) come from a 3-step process. The first step requires estimating Eq. 6 with all the base and geographic controls and city effects. In addition, racial composition changes are included as a control rather than an endogenous variable of interest. The highway variation is only used for estimating residential elasticity θ_r as the coefficient on $\Delta \log CMA_{igr}$ where the instruments for CMA changes are $CMA^{Plans/Rays}_{igr}$ as well as the CMA instruments in other neighborhoods for additional power $\{CMA^{Plans/Rays}_{igr,3-5}, CMA^{Plans/Rays}_{igr,5-10}, CMA^{Plans/Rays}_{igr,10-15}\}$. The elasticity estimates are presented in Table E.22, with values from Column 2 entering into the next step. Setting $\theta_N = 0.62, \theta_W = 0.75$ and taking the estimate of local costs from Brinkman and Lin (2022) where $b_{HW} = 0.175, \eta = 1.28$, I solve the quantitative model where endogenous amenities are removed. I simulate the construction of Interstate highways only for segments between 1960 and 1970. This counterfactual predicts racial composition changes under the assumption that other fundamentals of amenities and productivity are unchanged $\Delta \bar{b}_{igr} = 0$, $\Delta a_j = 0$. The prediction for racial composition $\widehat{L_{iW}/L_i}$ is used for the calculation of racial composition changes in the instrument $\Delta \log \widehat{(L_{iW}/L_i)} = \log(\widehat{L_{iW}/L_i}) - \log(L_{iW,1960}/L_{i,1960})$. The final set of instruments also includes the CMA instruments in other neighborhoods for additional power $\{CMA^{IV}_{igr,3-5}, CMA^{IV}_{igr,5-10}, CMA^{IV}_{igr,10-15}\}$.

CMA Instruments – Following that there are race-specific responses to CMA, the final set of instruments are CMA for each group separately. Variation in commute times again comes from either the planned routes or ray network for exogeneity. Specifically, the instruments are $\{CMA^{IV}_{iLB}, CMA^{IV}_{iHB}, CMA^{IV}_{iLW}, CMA^{IV}_{iHW}\}$.

E.2.7 Border Discontinuity

In Appendix Table E.24 Panel, I measure how various natural amenities change at the border. The variables of open water and wetlands, which are less manipulable by human intervention, are continuous, which supports the identification assumption of neighborhood features being similar. However, other variables, such as the amount of tree cover or the extent of urban development, are discontinuous. Still, these other factors may not be in fundamental amenities since they are correlated with prices, racial composition, or the SES controls, which are already residualized out, and the identification assumption is not necessarily violated.

In Appendix Table E.23, I take a reduced form approach to decomposing the discontinuity, reducing the reliance on exact parameter magnitudes. Instead of using the parameters to invert for amenities, I residualize population on prices and commuter access, racial composition, and demographic controls successively with redlining fixed effects, using the cross-sectional variation within neighborhood types, and project the residuals onto the border. This cross-sectional variation likely contains bias compared to the estimated parameters, but I report the results as a robustness check. The estimates are broadly the same where in Column 4, the discontinuity is still large for Black households at 0.489 (0.204) and continues to be zero for White households at 0.021 (0.079).

Lastly, I present additional results at the bottom of Appendix Table E.24. In Panel B, I probe the sensitivity of the drop in percent White across the redlined border by: (1) adding and removing the border fixed effects, (2) forming balanced samples where the number of tracts on the redlined and non-redlined sides is the same, and (3) altering the restrictiveness of how many neighborhoods are dropped away from physical barriers and school district borders. These adjustments do not greatly alter the findings. In Panel C, I further display the border discontinuity estimates for control variables used to residualize the fundamental amenities. Lastly, in Panel D,

Table E.22: Elasticity of Population to Commuter Market Access for Instruments

	(1)	(2)
	$\Delta \log L_{igr}$ (Δ Log Pop)	
Variables	Plans	Rays
$\Delta \log CMA_{igr}$		
Black	0.665*	0.624*
	(0.365)	(0.333)
	[0.587]	[0.550]
White	0.430**	0.745***
	(0.174)	(0.172)
	[0.578]	[0.628]
R-squared	0.517	0.513
Rounded Obs	58000	58000
C-D F-Stat	313.7	259.5
K-P F-stat	26.20	25.97

Notes: Data is tract by group for the first difference 1960-1970 using restricted Census microdata. CBSA fixed effects are interacted with race and education. Standard errors are cluster-robust by tract. Conley (1 km) standard errors in brackets. Controls are changes in log of rent, pct White, and 5 1-mile wide bins for distance from highways (built 1960-1970). Redlining fixed effects included. The geographic controls are log distance from CBD, rivers, lakes, shores, ports, historical railroads, canals, and historical large roads. All controls are interacted with race and education. Observation counts are rounded to the nearest 500 to meet Census disclosure rules. Borusyak and Hull (2023) control for CMA in large roads included. Kleibergen-Paap rk Wald and Cragg-Donald Wald F stats are reported. *** p<0.01, ** p<0.05, * p<0.1

Table E.23: Border Discontinuity Decomposition – Reduced Form Approach

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Black				*White*			
Variables	$\log L_{igr}$	Controls 1	Controls 2	Controls 3	$\log L_{igr}$	Controls 1	Controls 2	Controls 3
ψ_r: Border RD	1.425***	1.414***	0.555***	0.489**	-0.546***	-0.556***	-0.101	0.0205
	(0.226)	(0.227)	(0.212)	(0.204)	(0.122)	(0.122)	(0.0855)	(0.0794)
Bandwidth (mi)	0.495	0.488	0.414	0.365	0.358	0.364	0.380	0.397
Rounded Obs	13000	13000	13000	13000	13500	13500	13500	13500

Notes: Unit of observation is census tract by border in the redlining maps. Data comes from the 1960 restricted Census microdata. The dependent variable is residualized on fixed effects for education within race and on border fixed effects for all specifications. Controls 1 are log rent and log commuter access. Controls 2 includes Controls 1 and adds log percentage white. Controls 3 includes Controls 2 and adds log percentage high school grad, log population density, log average income, log percentage top quintile, log percentage bottom quintile, and log home values. Coefficients on controls are estimated with redlining fixed effects. Sample is limited to tracts that are at least 0.1 miles away from possible physical barriers such as historical large urban roads, constructed highways in 1960, or historical railroads and also at least 0.1 miles away from a school district boundary. The bandwidth is chosen optimally following Calonico et al. (2014). Distance from the border is measured in miles. Observation counts are rounded to the nearest 500 to meet Census disclosure rules. *** p<0.01, ** p<0.05, * p<0.1

I show how segregation along the border has changed over time from 1950 to 1990. The discontinuity in racial composition was largest in 1960 and 1970 and declined dramatically in 1980 after a decade of fair housing initiatives post Civil Rights legislation.

Table E.24: Border Discontinuity on Additional Variables

	(1)	(2)	(3)	(4)	(5)
	Panel A – Natural Amenities				
Variables	Pct Open Water	Pct Woody Wetlands	Pct Decid Forest	Pct Highly Developed	Pct Tree Cover
Border RD	0.005	-0.004	-0.015**	0.033***	-0.056***
	(0.003)	(0.003)	(0.005)	(0.007)	(0.009)
Dep. Var Mean	0.014	0.021	0.050	0.063	0.197
Bandwidth (mi)	0.406	0.315	0.351	0.347	0.461
Border FE	Yes	Yes	Yes	Yes	Yes
Observations	11529	11529	11529	11529	11529
	Panel B – Pct White in 1960				
Variables	Standard	Balanced Sample	Border FE	Drop Roads, Schools (0.1 mi)	Drop Roads, Schools (0.3 mi)
Border RD	-0.189***	-0.175***	-0.180***	-0.171***	-0.166***
	(0.027)	(0.029)	(0.025)	(0.026)	(0.034)
Bandwidth (mi)	0.351	0.368	0.355	0.403	0.432
Border FE	No	No	Yes	Yes	Yes
Observations	12573	5914	12532	10703	5717
	Panel C – Socioeconomic Variables in 1960				
Variables	Pct HS	Pct Bottom Q5	Pct Top Q5	Home Value	Rent
Border RD	-0.064***	0.104***	-0.093***	-22248***	-110.06***
	(0.006)	(0.010)	(0.008)	(3309)	(12.05)
Dep. Var Mean	0.265	0.189	0.207	114238 (2010$)	534 (2010$)
Bandwidth (mi)	0.267	0.397	0.409	0.428	0.248
Border FE	Yes	Yes	Yes	Yes	Yes
Observations	12275	12310	12310	12260	12268
	Panel D – Pct White Over Time				
Variables	1950	1960	1970	1980	1990
Border RD	-0.141***	-0.187***	-0.185***	-0.151***	-0.146***
	(0.023)	(0.029)	(0.035)	(0.037)	(0.035)
Dep. Var Mean	0.945	0.911	0.852	0.773	0.668
Bandwidth (mi)	0.350	0.350	0.350	0.350	0.350
Border FE	Yes	Yes	Yes	Yes	Yes
Observations	9964	9964	9964	9964	9964

Notes: Unit of observation is census tract by redlining border. Data comes from 1950, 1960, 1970, 1980, and 1990 tract-level aggregates retrieved from IPUMS NHGIS. The dependent variable is residualized on border fixed effects for many specifications. The balanced sample has the same number of tracts on the redlined and non-redlined sides. The "Drop Roads, Schools" sample is limited to tracts that are at least 0.1 (or 0.3) miles away from possible physical barriers such as historical large urban roads, constructed highways in 1960, or historical railroads and 0.1 (or 0.3) miles away from a school district boundary. The bandwidth is chosen optimally following Calonico et al. (2014) except for in Panel C, where the bandwidth is set to 0.35 so the effective sample remains the same across decades. The order of polynomial is 1 for all specifications. Distance from the border is measured in miles. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

F Welfare and Counterfactuals

F.1 Derivation of Direct Impacts to Welfare

This section derives the approximation of changes in welfare from total differentiating Eq. (7) with respect to the two variables that are changing due to the Interstate highway system: commute times t_{ijgr} and amenities B_{igr}. Assuming that commute times only affect commuter access and amenities do not affect any other indirect residential characteristics such as prices, the approximation is then

$$d\log U_{gr} = \sum_{i,j} \frac{\partial \log U_{gr}}{\partial t_{ijgr}} \Delta t_{ijgr} + \sum_{i} \frac{\partial \log U_{gr}}{\partial B_{igr}} \Delta B_{igr}$$

For ease of notation, define the location-specific utility shifter for neighborhoods, ignoring the idiosyncratic shock, as $V_{igr} = B_{igr} CMA_{igr} Q_i^{\beta_{gr}-1}$. Calculating the partial derivative for amenities first, the expression is as follows

$$\begin{aligned}
\frac{\partial \log U_{gr}}{\partial B_{igr}} &= \frac{1}{\theta_r} \frac{\partial V_{igr}^{\theta_r} / \partial B_{igr}}{\sum_s V_{sgr}^{\theta_r}} \\
&= \frac{V_{igr}^{\theta_r - 1}}{\sum_s V_{sgr}^{\theta_r}} CMA_{igr} Q_i^{\beta_{gr}-1} \\
&= \pi_{igr} / B_{igr}
\end{aligned}$$

where the last step substitutes in the residential share. A similar first step precedes calculating the partial derivative for commute times.

$$\begin{aligned}
\frac{\partial \log U_{gr}}{\partial t_{ijgr}} &= \frac{1}{\theta_r} \frac{\partial V_{igr}^{\theta_r} / \partial t_{ijgr}}{\sum_s V_{sgr}^{\theta_r}} \\
&= \frac{V_{igr}^{\theta_r - 1}}{\sum_s V_{sgr}^{\theta_r}} B_{igr} Q_i^{\beta_{gr}-1} \left(\partial CMA_{igr} / \partial t_{ijgr} \right) \\
&= \frac{V_{igr}^{\theta_r - 1}}{\sum_s V_{sgr}^{\theta_r}} B_{igr} Q_i^{\beta_{gr}-1} \left(-\Phi_{igr}^{\frac{1}{\phi}} \frac{T_{jgr} (w_{jgr} / d_{ijgr})^{\phi}}{\Phi_{igr}} \frac{\kappa_{gr}}{t_{ijgr}} \right) \\
&= -\frac{V_{igr}^{\theta_r}}{\sum_s V_{sgr}^{\theta_r}} \frac{T_{jgr} (w_{jgr} / d_{ijgr})^{\phi}}{\Phi_{igr}} \frac{\kappa_{gr}}{t_{ijgr}} \\
&= -\pi_{igr} \pi_{j|igr} \frac{\kappa_{gr}}{t_{ijgr}}
\end{aligned}$$

where the last step substitutes in the residential share and the conditional commuting share. Finally, note that $\Delta B_{igr} / B_{igr} = (-b_{highway} \exp(-\eta d_{i,highway}))$ so the direct impact to welfare is

$$d\log U_{gr} = -\sum_{i,j} \pi_{igr} \pi_{j|igr} \kappa_{gr} \Delta t_{ijgr} / t_{ijgr} - \sum_{i} \pi_{igr} b_{highway} \exp(-\eta d_{i,highway})$$

F.2 Solving for the Partial Equilibrium Counterfactual

To solve for the model counterfactuals, I employ a combination of observed data on travel times and city-level population $\{t_{ijgr}, \mathbb{L}_{gr}\}$, model parameters $\{\beta_{gr}, \kappa_{gr}, \phi, \theta_r, \mu\}$ with the externality parameter $\{\rho_r\}$, location fundamentals from the inversion process $\{b_{igr}\}$, and other location characteristics inferred during model inversion $\{k_i\}$. Equilibrium objects on the workplace side are fixed to their initial values for $\{w^0_{jgr}, \theta^0_i\}$. I assume starting values for the endogenous variables that correspond to the observed equilibrium for housing prices and the partially endogenous amenities $\{Q^0_i, B^0_{igr}\}$. From these starting values, I iterate following the equilibrium conditions of the model to reach a new equilibrium $\{Q^1_i, B^1_{igr}\}$.

$$\pi^1_{Rigr} = \frac{\left(B^0_{igr} CMA_{igr} (Q^0_i)^{\beta_{gr}-1}\right)^{\theta_r}}{\sum_t \left(B^0_{tgr} CMA_{tgr} (Q^0_{tr})^{\beta_{gr}-1}\right)^{\theta_r}} \quad \text{with } CMA_{igr} = \Phi_{igr}^{\frac{1}{\phi}}$$

$$\Phi_{igr} = \sum_s T_{sgr} (w^0_{sgr} / d_{isgr})^{\phi}$$

$$L^1_{Rigr} = \pi^1_{Rigr} \mathbb{L}_{gr}$$

$$Q^1_i = \left(\frac{Exp_i}{\theta^0_i k_i}\right)^{\mu} \quad \text{with } Exp_i = \sum_{g,r} (1-\beta_{gr}) \left(\sum_j \pi_{j|igr} w^0_{jgr}\right) L^1_{Rigr}$$

$$\text{and } \pi_{j|igr} = \frac{T_{jgr}(w^0_{jgr}/d_{ijgr})^{\phi}}{\sum_s T_{sgr}(w^0_{sgr}/d_{isgr})^{\phi}}$$

$$B^1_{igr} = b_{igr} (L^1_{RiW} / L^1_{Ri})^{\rho^R_r}$$

I continue the iterative procedure until the endogenous variables converge such that

$$\left\| \{Q^0_i, B^0_{igr}\} - \{Q^1_i, B^1_{igr}\} \right\| < \varepsilon$$

for some tolerance level ε. Before I reach that point, I update the endogenous variables as weighted averages of the initial values and the predicted values with $\lambda \in (0,1)$ following

$$Q^2_i = \lambda Q^1_i + (1-\lambda) Q^0_i$$
$$B^2_{igr} = \lambda B^1_{igr} + (1-\lambda) B^0_{igr}$$

F.3 Solving for the General Equilibrium Counterfactual

To solve for the model counterfactuals, I employ a combination of observed data on travel times and city-level population $\{t_{ijgr}, \mathbb{L}_{gr}\}$, model parameters $\{\beta_{gr}, \kappa_{gr}, \phi, \theta_r, \alpha, \alpha_{jg}, \alpha_{jgr}, \sigma^g,$
$\sigma^r, \mu\}$ with the externality parameters $\{\rho_r, \gamma^A\}$, location fundamentals from the inversion process $\{b_{igr}, a_i\}$, and other location characteristics inferred during model inversion $\{k_i, T_{jgr}\}$. I assume starting values for the endogenous variables that correspond to the observed equilibrium for wages, prices, distribution of rents, floorspace allocation and the partially endogenous amenities and productivity $\{w^0_{jgr}, Q^0_i, \theta^0_i, B^0_{igr}, A^0_i\}$. From these starting values,

I iterate following the equilibrium conditions of the model to reach a new equilibrium $\{w^1_{jgr}, Q^1_i, \theta^1_i, B^1_{igr}, A^1_j\}$.

$$\pi^1_{Rigr} = \frac{\left(B^0_{igr} CMA_{igr} (Q^0_i)^{\beta_{gr}-1}\right)^{\theta_r}}{\sum_t \left(B^0_{tgr} CMA_{tgr} (Q^0_{tr})^{\beta_{gr}-1}\right)^{\theta_r}} \quad \text{with } CMA_{igr} = \Phi_{igr}^{\frac{1}{\phi}}$$

$$\Phi_{igr} = \sum_s T_{sgr} (w^0_{sgr}/d_{isgr})^{\phi}$$

$$L^1_{Rigr} = \pi_{igr} \mathbb{L}_{gr}$$

$$\pi^1_{j|igr} = \frac{T_{jgr}(w^0_{jgr}/d_{ijgr})^{\phi}}{\sum_s T_{sgr}(w^0_{sgr}/d_{isgr})^{\phi}}$$

$$L^1_{Fjgr} = \sum_i \pi_{j|igr} L^1_{Rigr}$$

$$Y^1_i = A_i N_i^{\alpha} H_{Fi}^{1-\alpha} \quad \text{with } N_{jg} = \left(\sum_r \alpha_{jgr} (L^1_{Fjgr})^{\frac{\sigma^r-1}{\sigma^r}}\right)^{\frac{\sigma^r}{\sigma^r-1}} \quad N_j = \left(\sum_g \alpha_{jg} N_{jg}^{\frac{\sigma^g-1}{\sigma^g}}\right)^{\frac{\sigma^g}{\sigma^g-1}}$$

$$N_i = N_j / S_j \text{ for } i \in Tracts_j$$

$$H_{Fi} = (1-\theta^0_i) k_i (Q^0_i)^{\frac{1-\mu}{\mu}}$$

$$Q^1_i = \left(\frac{Exp_i + (1-\alpha)Y^1_i}{k_i}\right)^{\mu} \quad \text{with } Exp_i = \sum_{g,r} (1-\beta_{gr}) \left(\sum_j \pi^1_{j|igr} w^0_{jgr}\right) L^1_{Rigr}$$

$$\theta^1_i = \frac{Exp_i}{(Q^1_i)^{\frac{1}{\mu}} k_i}$$

$$w^1_{jgr} = (\alpha_{jgr}\omega_{jg}) \left(\frac{\alpha_{jg} W_j}{\omega_{jg}}\right)^{\frac{\sigma^g}{\sigma^r}} \left(\frac{\alpha Y^1_j}{W_j L^1_{Fjgr}}\right)^{\frac{1}{\sigma^r}} \text{ with } \omega_{jg} = \left(\sum_r \alpha_{jgr}^{\sigma^r} (w^0_{jgr})^{1-\sigma^r}\right)^{\frac{1}{1-\sigma^r}}$$

$$W_j = \left(\sum_g \alpha_{jg}^{\sigma^g} \omega_{jg}^{1-\sigma^g}\right)^{\frac{1}{1-\sigma^g}} \qquad Y^1_j = \sum_{i \in Tracts_j} Y^1_i$$

$$B^1_{igr} = b_{igr} (L^1_{RiW}/L^1_{Ri})^{\rho_r}$$

$$A^1_i = a_i (L^1_{Fj}/K_j)^{\gamma^A} \text{ for } i \in Tracts_j$$

I continue the iterative procedure until the endogenous variables converge such that

$$\left\| \{w^0_{jgr}, Q^0_i, \theta^0_i, B^0_{igr}, A^0_j\} - \{w^1_{jgr}, Q^1_i, \theta^1_i, B^1_{igr}, A^1_j\} \right\| < \varepsilon$$

for some tolerance level ε. Before I reach that point, I update the endogenous variables as weighted averages of the initial values and the predicted values with $\lambda \in (0,1)$ following

$$\begin{aligned}
w_{jgr}^2 &= \lambda w_{jgr}^1 + (1-\lambda) w_{jgr}^0 \\
Q_i^2 &= \lambda Q_i^1 + (1-\lambda) Q_i^0 \\
\theta_i^2 &= \lambda \theta_i^1 + (1-\lambda) \theta_i^0 \\
B_{igr}^2 &= \lambda B_{igr}^1 + (1-\lambda) B_{igr}^0 \\
A_j^2 &= \lambda A_j^1 + (1-\lambda) A_j^0
\end{aligned}$$

F.4 Sufficient Conditions for Uniqueness of Equilibria

The equilibrium defined has many sources of spillovers. The most immediate are through endogenous amenities and productivity from racial composition and agglomeration. Additional spillovers emerge through inelastic land generating a congestion force in housing supply and the idiosyncratic preferences of individuals creating dispersion forces. As the wages of each group depend on the labor supply of other workers, there are productivity spillovers across groups at workplaces.

I follow Allen and Arkolakis (2022) where I rewrite the equilibrium conditions as a set of H types of economic interactions conducted by the set of N heterogeneous agents. I then construct the $H \times H$ matrix of the uniform bounds of the elasticities on the strength of economic interactions. The equilibrium system falls under a constant elasticity form that is commonly used in spatial economics. Building on Tsivanidis (2023), I reformulate the CMA measures as solutions to a system of equations in residential and workplace populations and commute costs. With these conditions on model parameters, I derive theory-consistent equations to estimate parameter values in the next section.

First, I rewrite the equilibrium conditions in a form that adheres to the constant elasticity system of Allen and Arkolakis (2022) where spillovers are of an exponential form. I further allow the elasticities to differ by the type of the agent.

$$x_{ih} = f_{ijh}(x_j) = \sum_{j \in \mathcal{N}} K_{ijh} \prod_{h' \in \mathcal{H}} x_{jh'}^{\alpha_{ihh'}}$$

In this setting, type $\mathcal{N}$ can be a combination of location $i \in \{1, \dots, S\}$, education $g \in \{L, H\}$, and race $r \in \{B, W\}$. The set of economic interactions $\mathcal{H}$ include population, prices, amenities, and productivity. Define the city-level constant following $\lambda_{gr} = \mathbb{L}_{gr} U_{gr}^{-\theta_r}$. The equilibrium conditions are the stacked set of equations

$$\begin{aligned}
L_{igr} &= \lambda_{gr} \left(B_{igr} Q_i^{\beta_{gr}-1} \Phi_{igr}^{\frac{1}{\phi}} \right)^{\theta_r} \\
L_{Fjgr} &= \lambda_{gr} T_{jgr} w_{jgr}^{\phi} \Phi_{Fjgr} \\
\Phi_{igr} &= \lambda_{gr} \sum_j d_{ijgr}^{-\phi} \frac{L_{Fjgr}}{\Phi_{Fjgr}} \\
\Phi_{Fjgr} &= \lambda_{gr} \sum_i d_{ijgr}^{-\phi} \frac{L_{igr}}{\Phi_{igr}} \\
N_i^{(\sigma_g-1)/\sigma_g} &= \sum_g \alpha_{ig} N_{ig}^{(\sigma_g-1)/\sigma_g} \\
N_{ig}^{(\sigma_r-1)/\sigma_r} &= \sum_r \alpha_{igr} L_{Figr}^{(\sigma_r-1)/\sigma_r}
\end{aligned}$$

$$Y_i = A_i^{1/\alpha} N_i (1-\alpha)^{(1-\alpha)/\alpha} Q_i^{(\alpha-1)/\alpha}$$

$$\bar{w}_{igr} = \sum_j T_{jgr} w_{jgr}^{\phi+1} d_{ijgr}^{-\phi} \Phi_{igr}^{-1}$$

$$Q_i^{\frac{1}{\mu}} = k_i^{-1} \left(\sum_{g,r} (1-\beta_{gr}) \bar{w}_{igr} L_{igr} + (1-\alpha) Y_i \right)$$

$$w_{jgr} = \alpha_{jgr} (\omega_{jg})^{\frac{\sigma_r - \sigma_g}{\sigma_r}} \alpha_{jg}^{\frac{\sigma^g}{\sigma^r}} W_j^{\frac{\sigma^g - 1}{\sigma^r}} (\alpha Y_j)^{\frac{1}{\sigma^r}} L_{F\,jgr}^{-\frac{1}{\sigma_r}}$$

$$\omega_{jg}^{1-\sigma^r} = \sum_r \alpha_{jgr}^{\sigma^r} w_{jgr}^{1-\sigma^r}$$

$$W_j^{1-\sigma^g} = \sum_g \alpha_{jg}^{\sigma^g} \omega_{jg}^{1-\sigma^g}$$

$$L_{iW} = \sum_g L_{igW}$$

$$L_i = \sum_{g,r} L_{igr}$$

$$B_{igr} = b_{igr} L_{iW}^{\rho_r} L_i^{-\rho_r}$$

$$A_j^{1/\gamma^A} = \frac{a_j^{1/\gamma^A}}{K_j} \sum_{g,r} L_{F\,jgr}$$

Almost all of the elasticities $\varepsilon_{ijh,jh'}(x_j) = \frac{\partial \log f_{ijh}(x_j)}{\partial \log x_{jh'}}$ are of the form $\varepsilon_{ijh,jh'}(x_j) = \alpha_{hh'}$ except for the elasticity to price and the spillovers of racial composition on residential location choices. Let $\beta = \min_{g,r}\{\beta_{gr}\}$ where $\beta_{g,r} > 0$, $\rho = \max_r\{|\rho_r|\}$, and $\theta = \max_r\{\theta_r\}$. Then the $H \times H$ matrix $(\mathbf{A})_{hh'}$ where $H = 16$ is

$$\begin{bmatrix}
0 & 0 & \theta/\phi & 0 & 0 & 0 & 0 & 0 & \mu(\beta-1)\theta & 0 & 0 & 0 & 0 & 0 & \theta & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & \phi & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \frac{(\sigma_g-1)\sigma_r}{(\sigma_r-1)\sigma_g} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & (\frac{\sigma_r-1}{\sigma_r}) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & (\frac{\sigma_g}{\sigma_g-1}) & 0 & 0 & 0 & \mu(\frac{\alpha-1}{\alpha}) & 0 & 0 & 0 & 0 & 0 & 0 & \frac{\gamma^A}{\alpha} \\
0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & \phi+1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & -\frac{1}{\sigma_r} & 0 & 0 & 0 & 0 & \frac{1}{\sigma_r} & 0 & 0 & 0 & \frac{\sigma_r-\sigma_g}{\sigma_r(1-\sigma_r)} & -\frac{1}{\sigma_r} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1-\sigma_r & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1-\sigma_g}{1-\sigma_r} & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \rho^R & -\rho^R & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}$$

Following Allen and Arkolakis (2022) Theorem 1 Part (i), a sufficent condition for uniqueness is that the spectral radius $\rho(A) < 1$. At the current parameter values in Table 4, $\rho(A) > 1$. However, unique equilibria may exist with the listed parameters as the above condition is sufficient but not necessary for uniqueness.

G Data

Decennial Census Microdata

The Decennial Census is the data source for all of the quantitative estimation. Residential population, workplace population, commute flows, rental prices, and other characteristics of locations come from the Decennial Census microdata in 1960 and 1970. The decade between these two censuses covers a substantial portion of highway construction as by 1960, 20% of the national network was completed, and by 1970, 71% was completed.

- **Residences** – Residential geographic units are census tracts that represent neighborhoods with the usual tract containing 2,500 to 8,000 people. In 1960, there were 42,689 tracts across the United States but not all of the country was contained within tracts. By 1990, the entire country fell under some tract definition, and in 2010, there were 73,175 tracts. Tracts are re-defined across census surveys as population levels across neighborhoods change, so I interpolate all tract-level aggregates to consistent tract definitions with the Longitudinal Tract Database to match 2010 Census delineations (see more below) (Logan et al., 2014). The shapefile for 2010 census tract definitions is retrieved from IPUMS NHGIS.

- **Workplaces** – I construct geographic units which I define as Place of Work (POW) Zones from the Journey to Work questions of the 1960 Census, the first survey in which the Census Bureau asked for location of employment. County and municipality of place of work are reported as 1960-specific Universal Area Codes (UAC), and from these UACs, I calculate the smallest intersection of county and municipality to create the POW Zone. These POW Zones are then overlayed on 1960 tract definitions to create a spatial unit and mapped into 2010 tract boundaries with a crosswalk between 1960 and 2010 tracts. As the UAC for place of work is missing for some observations, I reweight the microdata by calculating inverse probability weights based on observed demographic variables of age, age squared, educational attainment, employment status, total income, wages, industry, occupation, a poverty indicator, race, gender, mode of transport, weeks worked, and a urban/rural indicator. In 1970, place of work is available for UACs, although 1970 UACs are different units from 1960 UACs. For some observations, place of work at the tract level is observed. The inverse probability weights for the 1970 Census are based on whether UAC is observed. For those with tract-level place of work, I assign them to the tract. For those with only UAC, I evenly distribute them across the tracts that are in the UAC. The 1970 tract reweighted sums are then mapped into the 1960 POW zones using a crosswalk between 1970 tracts and 2010 tracts to create a panel of workplace data from 1960 to 1970.

- **Cities** – Cities are represented by Metropolitan Statistical Areas out of the Core-Based Statistical Areas (CBSAs) from 2010 Census definitions. The quantitative analysis requires granular data on commute flows to workplaces from the Decennial microdata in 1960 and 1970. To create the POW Zone, the sample of cities is smaller. While some cities have many unique counties and municipalities, others have very few. For there to be sufficient spatial granularity in place of work, I limit the sample of cities to 25 of the largest, and these cities in total contain 406 POW Zones. I provide the list of cities with available data in Appendix Table G.26. For the motivating empirical analysis, the sample of cities is limited to the 100 cities with Yellow Book maps using public-use tract-level aggregates from NHGIS (see below).

- **Commuting** – With residences as tracts and workplaces as POW zones, commute flows are constructed from population counts over the cross-product of residences and workplaces and are comparable to the widely used Census Transportation Planning Package (CTPP) for commuting after 1990. Starting with the 1980 census, commute times are reported in the Journey to Work section. While individuals may be using non-automobile modes of transport during this time, the lack of data on public transit across a large set of cities makes analysis of other modes difficult. I account for commuting through other methods by assuming public transit systems and walking have not changed in speed from 1960 to 1980 and take reported commute times from the 1980 Decennial Census microdata, the first census survey with commute time data. I non-parametrically estimate non-automobile commute times over 15 bins of Euclidean distance for bilateral pairs of tract of residence and POW Zone and 3 bins of distance from the CBD for both residences and workplaces. The 15 bins of Euclidean distance are fully interacted with the 3 residential bins and 3 workplace bins. Adjusting for distance from the central city captures how car usage is greater when workers live in the suburbs or commute to the suburbs for employment. For each race and education group, I similarly create mode of transport weights over the interaction of bins of Euclidean distance in 1960 and 1970 and bins of distance from the CBD for residences and workplaces. Weighted commute times are averages using the weight for automobile modes (private auto, carpool, van or truck) with the computer generated commuting times for the road network (see below) and the weight for non-automobile modes with the binned commute times from above.

- **Race by Education Tabulations –** To tabulate the population counts, the Census Long-Form person-level sample (25% in 1960 and 15% in 1970) is limited to workers and divided into race and education categories. Person-level sampling weights from the Census are used for all tabulations. Race is divided into White and Non-White as finer splits of race leave too few counts for smaller geographic units. Education is also separated into two categories where those with a high school degree or higher are considered highly-educated and those without a high school degree are considered less-educated. Wages are then calculated for each geographic unit by race and education.

- **Housing Prices –** Housing price data come from the household-level sample with household sampling weights used for all tabulations. Quality-adjusted rents per unit are calculated by taking the rent and residualizing out housing characteristics of the number of rooms, bedrooms, and bathrooms,the availability of a basement, kitchen, heat, hot water, shower/bathtub, indoor toilet, and the year built (setting as the base price the average over the fitted values of housing characteristics for the CBSA and then adding in the neighborhood fixed effects for each neighborhood).

Digitized Roads and Highway Routes

- **Historical Urban Roads –** To capture commuting on the road network prior to Interstate construction, I digitize maps of historical U.S. and state highways and major roads from Shell Atlases in 1951 and 1956 (Rumsey, 2020). To create maps of the historical roads, I start with a highly accurate digital map of modern day major roads from ESRI (2019). I remove Interstate highways and keep major roads less than a freeway, other major roads, and secondary roads as a starting point for the historical map roads. I then georeference the Shell map images in ArcMap and edit the modern day major roads file to match the historical roads maps. I categorize the historical roads into two groups: Superhighways and other major roads following the legend of the Shell Atlases. Maps from 71 cities were digitized as shown in Appendix Table G.26. Examples for Baltimore, MA and San Francisco, CA are shown in Figure G.7.

- **Yellow Book Plans –** I retrieve maps of the planned routes from the *General Location of National System of Interstate Highways Including All Additional Routes at Urban Areas Designated in September 1955*, commonly known as the Yellow Book, for plans of Interstate highways within cities. While maps for 100 cities are available, some cities are located within the same CBSA (e.g. Dallas and Fort Worth) and some are Micropolitan Statistical Areas. For these reasons, in Appendix Table G.26, only 96 cities are shown. These planned maps were originally used by Brinkman and Lin (2022), and I manually digitized them for this project in ArcMap by georeferencing the map images and creating the spatial lines. Examples for Atlanta, GA and Cleveland, OH are depicted in Figure G.8.

- **National 1947 Plan –** I digitize a map of the 1947 plan of national highway routes from Baum-Snow (2007). This map has less spatial granularity compared to the Yellow Book plans but conveys the direction of routes between cities and which cities the Interstate system was designed to connect. The 1947 plan and Yellow Book maps are consolidated into one planned network.

- **Euclidean Rays** – I construct an additional network of highway routes following the planned routes where I connect cities and towns in the planned maps with straight line rays. This network follows the "inconsequential units" approach where neighborhoods that happen to be located between major cities are treated by the Interstate highway system.
- **Constructed Highways** – The constructed Interstate system comes from MIT Libraries' file of Interstate Highways in 1996 (ESRI, 1996). I exploit the panel variation in when different segments were built by combining this constructed network map with the PR-511 database on dates of construction from Baum-Snow (2007) to examine changes only on routes constructed between 1960 and 1970.

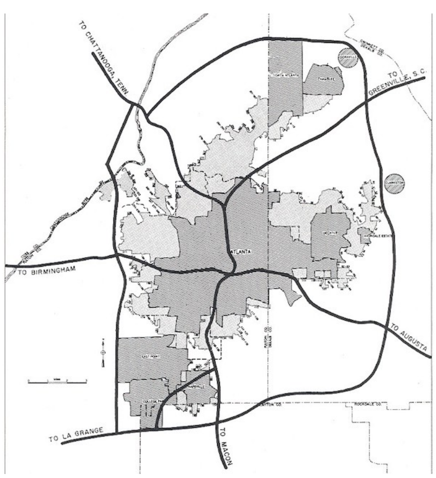
Figure G.7: Historical Roads from Shell Atlases for Baltimore and San Francisco

(a) Baltimore, Maryland

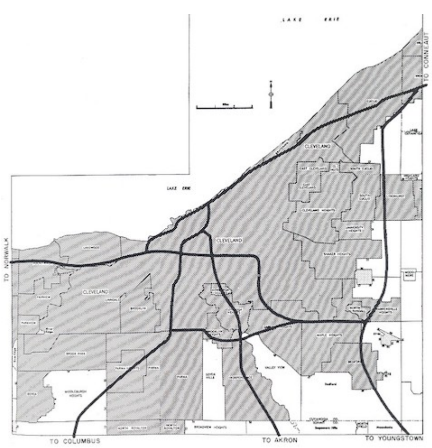
(b) San Francisco, California

Notes: Shell Atlases by the H.M. Gousha Company in 1956 retrieved from the Rumsey Collection for Baltimore and San Francisco.

Figure G.8: Yellow Book Maps for Atlanta and Cleveland



(a) Atlanta, Georgia



(b) Cleveland, Ohio

Notes: Yellow Book (General Location of National System of Interstate Highways Including All Additional Routes at Urban Areas Designated in September 1955) maps retrieved from the Bureau of Public Roads.

Commuting Networks

- **Speeds –** To calculate commute times on the road networks, I assume speeds for different segments of the routes. For the historical urban roads, large roads (superhighways) are set to have a speed of 40 mph while other major roads are set to have a speed limit of 30 mph following travel surveys conducted during the 1950-1960 period (Gibbons and Proctor, 1954; Walters, 1961). For constructed highways, I use the speed limit for each segment of the highway. The consolidated planned routes of the 1947 plan and Yellow Book maps do not have associated speed limits, so I assign each 2500 meter segment the speed limit of the nearest constructed highway. The Euclidean ray spanning network is set to 60 mph. Minor errors in assignment of speed limits should not affect the results too much given that for urban highways, speed limits cover a narrow range of 55 mph to 65 mph.

- **Commuting Matrices –** For the period prior to highway construction, I calculate commuting times from each 2010 delineated tract centroid to other tract centroids within the same CBSA using ArcNetwork Analyst. The only road network that is traversable is the major roads from the historical road maps. For the period during highway construction, I retrieve the highway network at two stages mid-construction: for all routes built before 1960 and for all routes built before 1970. I overlay these semi-completed highway networks on the historical road network to calculate commuting times during these intermediate periods to align with the years when data is available from the Decennial Census. Using the planned maps and Euclidean ray networks, I construct commute times for the instruments by overlaying the planned and ray networks instead of the Interstate routes on the historical road network. Since there is some distance from tract centroids to the nearest road, and ArcGIS sets the starting point as the point on the traversable network that is closest to the centroid, I add in the additional travel time from the centroid to the road assuming a travel speed of 20 mph. Least cost travel times between tracts are then generated following Dijkstra's algorithm in ArcGIS Network Analyst for 49 million pairwise comparisons. I validate that the computer generated commute times for the fully constructed highway network overlayed on the historical road network are closely correlated with reported commute times by automobile in the 1980 Census (despite possible further road development) in Appendix Table G.25. The 1980 Census is the first census survey with commute time data.

Table G.25: Commuting Time Comparison in 1980

Variables	Reported 1980 Commute Time (Minute)
Generated Commute Time (Minute)	0.683***
	(0.0122)
Constant	10.52***
	(0.395)
R-squared	0.537
Correlation Coefficient	0.733
Rounded Obs	11500

Notes: Unit of observation is Place of Work Zone by Place of Work Zone. Data comes from the 1980 Census for survey reported commute times of workers whose mode of transport is private automobile. Computer generated commute times use the full constructed highway network and historical urban roads. Observation counts are rounded to the nearest 500 to meet Census disclosure rules. Robust standard errors are included in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Geographic Features

- **Historical Rail, Canals, Rivers** – Historical railroad networks from 1826-1911, 19th century canals, and 19th century steam-boat navigated rivers are included as controls (Atack, 2015, 2016, 2017).

- **Natural Features** – Distances to natural features including lakes, shores, and ports come from Lee and Lin (2017) and are included as controls.

- **Central Business Districts** – Centroids of the central business districts of MSAs come from Holian and Kahn (2015) although their list does not cover the full list of cities studies in this paper. To obtain the location of other central business districts, I search for where central business districts are in the modern day (assuming most downtowns do not change their location) for cities in Google Earth.

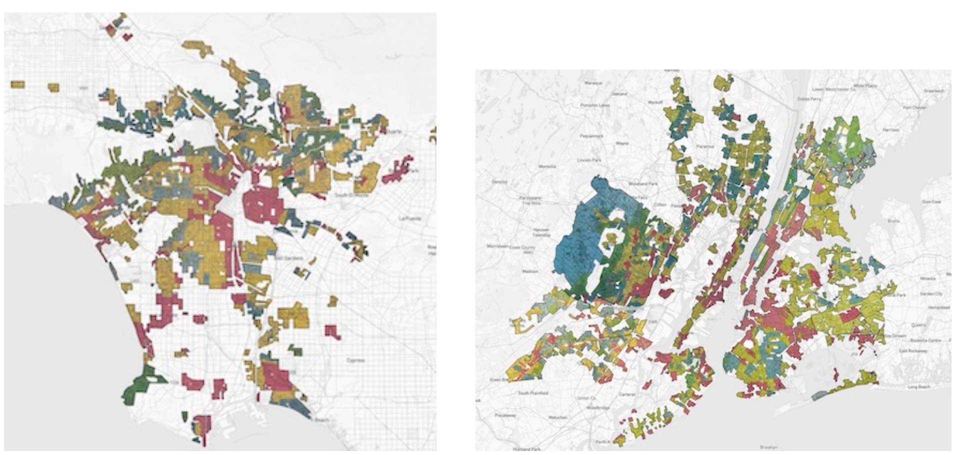
- **HOLC Redlining** – Redlining maps for the Home Owners' Loan Corporation come from a group of digital historians at Mapping Inequality: Redlining in New Deal America (Nelson et al., 2020). Examples for Los Angeles, CA and New York City, NY are in Figure G.9.

  - **Borders** – To calculate distances from tracts to borders in the HOLC maps for the border discontinuity, I find for each tract the distance to all HOLC map borders. I keep all borders that are within 2 km of the tract centroid. If the tract is redlined, then it has a positive distance from the redlining border. If the tract is non-redlined, the distance is negative.
  - **Redlined/Non-Redlined** – I calculate the percentage of each census tract that is redlined by overlaying the 2010 tract boundaries on the HOLC redlining maps. Tracts that are more than 80% covered by HOLC grade D areas are considered redlined. The results are not sensitive to the percentage cut-off as 70 percent of tracts are either 100% or 0% graded D. At the 80% cutoff, 3050 out of 13436 tracts are redlined while at a 50% cutoff, 3761 tracts are redlined.

- **School District Borders** – School district boundaries used for the border design are acquired from the National Center for Education Statistics (NCES) for the 1989-1990 school year.

- **Distances to Features** – I calculate the distance from tract centroids to each of the geographic features above. For the POW zones, I take the average of the distances from tract centroids for the tracts within a POW zone.

Figure G.9: Redlining Maps for Los Angeles and New York City



(a) Los Angeles, California (b) New York City, New York

Notes: HOLC Maps for Los Angeles and New York City from Mapping Inequality: Redlining in New Deal America.

Natural Amenities

- **Land Cover** – The National Land Cover Database (NLCD) from the U.S. Geological Survey provides data nationwide on land cover types at high spatial resolution (30m). I obtain the dataset for 2011 and limit the characteristics to the land cover types of open water, woody wetlands, developed high-intensity, and deciduous forest. Other land cover types that are available include barren land, cultivated crops, and perennial ice and snow.

- **Tree Canopy Cover** – The U.S. Forest Service Science provides a dataset on Tree Canopy Cover (TCC), and I obtain the 2011 version. It is a 30m spatial resolution file with one variable representing the percentage of canopy cover.

- **Overlap of Tracts with Natural Amenities** – I calculate the overlap between each 30m square from the NLCD and TCC datasets with each tract from the census tracts (with 2010 boundaries) shapefile. A weighted average is computed across the squares that overlap with census tracts.

Air Pollution Index

- **Environmental Pollution** – The Centers for Disease Control and Prevention (CDC) National Environmental Public Health Tracking Network generates air quality measures at the census tract-level using the Environmental Protection Agency (EPA)'s Downscaler model for the mean predicted concentration of PM 2.5. I obtain the 2001-2005 daily estimates and aggregate over the 5 years of data to create a tract-level average.

IPUMS NHGIS Public-Use Aggregates

- I construct a panel of tract-level characteristics from the public-use aggregates available at IPUMS NHGIS (Manson et al., 2017) starting from 1950 and ending in 2010. Aggregates include tract-level population by education, race, income, and housing rents and home values. This dataset is interpolated to be consistent with 2010 tract definitions and spans the full set of cities with planned (Yellow Book) maps in the U.S. The panel is unbalanced however as it was not until 1990 that the Census defined tract geographic units for the entire United States.

Longitudinal Tract Crosswalks

- Tract cross-walk weights derived using population overlaps from the Longitudinal Tract Database are available for 1970 to 2010 from Logan et al. (2014) to harmonize tract-level data across decades to 2010 boundaries. Weights for 1950 and 1960 come from Lee and Lin (2017) and are derived from area overlaps.

Table G.26: Data Availability by Metro Area

Metropolitan Statistical Area	Yellow Book	HOLC	Historical Roads	Census
Albany-Schenectady-Troy, NY	X	X	X	X
Allentown-Bethlehem-Easton, PA-NJ	X	X	X	X
Atlanta-Sandy Springs-Marietta, GA	X	X	X	X
Baltimore-Towson, MD	X	X	X	
Bangor, ME	X			
Baton Rouge, LA	X		X	
Battle Creek, MI	X	X	X	
Birmingham-Hoover, AL	X	X	X	
Boston-Cambridge-Quincy, MA-NH	X	X	X	X
Buffalo-Niagara Falls, NY	X	X	X	
Burlington-South Burlington, VT	X			
Chattanooga, TN-GA	X	X	X	
Chicago-Joliet-Naperville, IL-IN-WI	X	X	X	X
Cincinnati-Middletown, OH-KY-IN	X	X	X	
Cleveland-Elyria-Mentor, OH	X	X	X	X
Columbia, SC	X	X	X	
Columbus, OH	X	X	X	
Dallas-Fort Worth-Arlington, TX	X	X	X	X
Davenport-Moline-Rock Island, IA-IL	X	X		
Denver-Aurora-Broomfield, CO	X	X		
Des Moines-West Des Moines, IA	X	X		
Detroit-Warren-Livonia, MI	X	X	X	X
Erie, PA	X	X	X	
Eugene-Springfield, OR	X			
Flint, MI	X	X	X	
Fort Smith, AR-OK	X		X	
Gadsden, AL	X		X	
Grand Rapids-Wyoming, MI	X	X	X	
Great Falls, MT	X			
Greenville-Mauldin-Easley, SC	X		X	
Harrisburg-Carlisle, PA	X	X	X	
Hartford-West Hartford-East Hartford, CT	X	X	X	X
Houston-Sugar Land-Baytown, TX	X	X	X	X
Indianapolis-Carmel, IN	X	X	X	
Jackson, MS	X	X	X	
Kansas City, MO-KS	X	X	X	X
Kingsport-Bristol-Bristol, TN-VA	X			
Kingston, NY	X		X	
Knoxville, TN	X	X	X	
Lake Charles, LA	X		X	
Lansing-East Lansing, MI	X	X	X	
Lincoln, NE	X	X	X	
Little Rock-North Little Rock-Conway, AR	X	X	X	
Los Angeles-Long Beach-Santa Ana, CA	X	X	X	X
Louisville/Jefferson County, KY-IN	X	X	X	
Macon, GA	X	X		
Manchester-Nashua, NH	X	X		

Notes: The table displays 96 CBSAs because while there are 100 cities in the Yellow Book, not all of them have an associated Metropolitan Statistical Area. Some are in Micropolitan Statistical Areas, and two cities (Dallas and Fort Worth) are combined into one MSA. The HOLC redlining maps are available for more cities, but the table is restricted to the sample of Yellow Book maps. Historical road maps are also available for more cities, but only 71 are digitized in this paper. The Census column indicates which cities are included in the quantitative analysis using Decennial microdata.

Table G.26: Data Availability by Metro Area CONTINUED

Metropolitan Statistical Area	Yellow Book	HOLC	Historical Roads	Census
Memphis, TN-MS-AR	X	X	X	
Miami-Fort Lauderdale-Pompano Beach, FL	X	X	X	
Milwaukee-Waukesha-West Allis, WI	X	X	X	
Minneapolis-St. Paul-Bloomington, MN-WI	X	X	X	X
Monroe, LA	X		X	
Montgomery, AL	X	X	X	
Nashville-Davidson–Murfreesboro–Franklin, TN	X	X	X	
New Orleans-Metairie-Kenner, LA	X	X	X	
New York-New Jersey-Long Island, NY-NJ-PA	X	X	X	X
Oklahoma City, OK	X	X	X	
Omaha-Council Bluffs, NE-IA	X	X	X	
Pensacola-Ferry Pass-Brent, FL	X			
Peoria, IL	X	X	X	
Philadelphia-Camden, PA-NJ-DE-MD	X	X	X	X
Phoenix-Mesa-Glendale, AZ	X	X		
Pittsburgh, PA	X	X	X	X
Pocatello, ID	X			
Portland-South Portland-Biddeford, ME	X			
Portland-Vancouver-Hillsboro, OR-WA	X	X	X	X
Providence-New Bedford-Fall River, RI-MA	X	X	X	X
Rapid City, SD	X			
Reading, PA	X		X	X
Richmond, VA	X	X		
Roanoke, VA	X	X		
Rochester, NY	X	X	X	
Saginaw-Saginaw Township North, MI	X	X	X	
St. Joseph, MO-KS	X	X		
St. Louis, MO-IL	X	X	X	X
Salem, OR	X			
San Antonio-New Braunfels, TX	X	X		
San Francisco-Oakland-Fremont, CA	X	X	X	X
Seattle-Tacoma-Bellevue, WA	X	X		
Shreveport-Bossier City, LA	X	X	X	
Sioux Falls, SD	X			
Spartanburg, SC	X		X	
Springfield, MA	X	X	X	X
Syracuse, NY	X	X	X	
Tampa-St. Petersburg-Clearwater, FL	X	X	X	
Toledo, OH	X	X	X	
Topeka, KS	X	X		
Tucson, AZ	X			
Tulsa, OK	X	X	X	
Tuscaloosa, AL	X		X	
Utica-Rome, NY	X	X	X	
Virginia Beach-Norfolk-Newport News, VA-NC	X	X	X	X
Washington-Arlington, DC-VA-MD-WV	X		X	X
Wheeling, WV-OH	X	X	X	
Wichita, KS	X	X		
Worcester, MA	X		X	X

Notes: The table displays 96 CBSAs because while there are 100 cities in the Yellow Book, not all of them have an associated Metropolitan Statistical Area. Some are in Micropolitan Statistical Areas, and two cities (Dallas and Fort Worth) are combined into one MSA. The HOLC redlining maps are available for more cities, but the table is restricted to the sample of Yellow Book maps. Historical road maps are also available for more cities, but only 71 are digitized in this paper. The Census column indicates which cities are included in the quantitative analysis using Decennial microdata.