The problem has existed over endless years: Racialized difference in commuting, 1980–2019

devin michelle bunten
MIT (DUSP)

Ellen Fu
Penn (Wharton)

Lyndsey Rolheiser
UConn

Christopher Severen
FRB Philadelphia

March '22

Disclaimer: *This presentation represents preliminary research that is being circulated for discussion purposes. The views expressed in this paper are solely those of the authors and do not necessarily reflect those of the Federal Reserve Bank of Philadelphia or the Federal Reserve System. Any errors or omissions are the responsibility of the authors.* Nassir Holden and Nathan Schor have provided excellent research assistance.

“The problem has existed over endless years”

- Plessy v. Ferguson (1896), which legitimized doctrine of ‘separate but equal’, was about segregation on trains
- Quote from Dr. Martin Luther King Jr. about discrimination faced by Black bus riders, made during the Montgomery Bus Boycott (1955)
- LA Bus Riders Union vs. LA MTA (1990s) about bus vs. rail service quality

Photo of LA BRU supporters from
https://www.impactfund.org/social-justice-blog/bus-riders

$\Rightarrow$ Racialized difference in transportation is a pervasive component of US history

“The problem has existed over endless years”

Are commuting outcomes in American cities today equitable by race?

How has racialized difference in commutes evolved over the last 40 years?

This paper: Comprehensive accounting of racialized difference in commuting in the US

- Update prior literature in economics and sociology, study trend 1980–2019
- Primarily positive (rather than normative) analysis
- Consider role of both individual and aggregate (city-level) factors
- Suggest an interpretation related to the stratification of urban space

The Short Answer: **No**

...despite some improvement

1. Black commuters today commute 22.4 minutes/week more than White commuters
   - Difference declined from 50.3 min/wk since 1980, but still persists (even conditionally)
   - Amongst transit users, the difference has not decreased

The Short Answer: **No**

...despite some improvement

1. Black commuters today commute 22.4 minutes/week more than White commuters
   - Difference declined from 50.3 min/wk since 1980, but still persists (even conditionally)
   - Amongst transit users, the difference has not decreased
2. Mode and city of residence play key roles in explaining difference
   - From 24% (1980) to 50% (today) of difference explained by city, but none of the decline
   - Mode explains ~ 30% of difference and 24% of decline (mostly auto adoption)

The Short Answer: **No**

...despite some improvement

1. Black commuters today commute 22.4 minutes/week more than White commuters
   - Difference declined from 50.3 min/wk since 1980, but still persists (even conditionally)
   - Amongst transit users, the difference has not decreased
2. Mode and city of residence play key roles in explaining difference
   - From 24% (1980) to 50% (today) of difference explained by city, but none of the decline
   - Mode explains $\sim$ 30% of difference and 24% of decline (mostly auto adoption)
3. 40-50% of difference each year and 63% of decline is not explained by observables
   - Within-city res. location (PUMA) does not account for much of the difference
   - Difference largest at lower incomes, but are present at high incomes too

The Short Answer: **No**

...despite some improvement

1. Black commuters today commute 22.4 minutes/week more than White commuters
   - Difference declined from 50.3 min/wk since 1980, but still persists (even conditionally)
   - Amongst transit users, the difference has not decreased
2. Mode and city of residence play key roles in explaining difference
   - From 24% (1980) to 50% (today) of difference explained by city, but none of the decline
   - Mode explains ~ 30% of difference and 24% of decline (mostly auto adoption)
3. 40-50% of difference each year and 63% of decline is not explained by observables
   - Within-city res. location (PUMA) does not account for much of the difference
   - Difference largest at lower incomes, but are present at high incomes too
4. Differences persist mainly in large, segregated, congested, and *expensive* cities
   - City-specific estimates of difference correlate with these measures
   - Housing price IV and tighter correlation between n'hood price and travel time indicate spatial stratification
   - If housing $$ today were at 1980 levels, commute time difference would be 40% smaller

Literature & Contributions

- Document and quantify racialized difference in US commuting
  - Some work in sociology and econ documenting differences in travel times (Gabriel & Rosenthal 1996; Johnston-Anumonwo 1997, 2001; McLafferty 1997; Petitte & Ross 1999; Taylor & Ong 1995)
  - Related literature showing auto access important in US for URM LFP (Gautier & Zenou 2010; Gobillon, Selod, & Zenou 2007; Kain 1968; Ong 2002; Raphael & Stoll 2001; Ong & Miller 2005)

Literature & Contributions

- Document and quantify racialized difference in US commuting
  - **Contribution:** Update results and provide exhaustive coverage across US, focusing on commute times

Literature & Contributions

- Document and quantify racialized difference in US commuting
  - **Contribution:** Update results and provide exhaustive coverage across US, focusing on commute times
- Decomposition of difference by demographics
  - Extensive literature in gender and race wage differences (Altonji & Blank 1999; Blau & Kahn 2017; Chamberlain 2016; DiNardo, Fortin, & Lemieux 1995; Kitagawa 1955)
  - Regression-friendly approaches to decompositions (appropriate for large data) (Fortin 2008; Fortin, Lemieux, & Firpo 2011; Gelbach 2016)

Literature & Contributions

- Document and quantify racialized difference in US commuting
  - **Contribution:** Update results and provide exhaustive coverage across US, focusing on commute times
- Decomposition of difference by demographics
  - **Contribution:** Study new outcome and extend decomposition methods to study differences across locations

Literature & Contributions

- Document and quantify racialized difference in US commuting
  - **Contribution:** Update results and provide exhaustive coverage across US, focusing on commute times
- Decomposition of difference by demographics
  - **Contribution:** Study new outcome and extend decomposition methods to study differences across locations
- Spatial stratification and urban form
  - Transportation technology impacts urban form (Bento et al. 2005; Glaeser, Kahn, & Rappaport 2008; Heblich, Redding, & Sturm 2020; LeRoy & Sonstelie 1983)
  - Increased stratification (or pressure therefore) lately? (Guerrieri, Hartley, & Hurst 2013; Gyourko, Mayer, & Sinai 2013; Lee & Lin 2018; Van Nieuwerburgh & Weill 2010; Su 2021)

Literature & Contributions

- Document and quantify racialized difference in US commuting
  - **Contribution:** Update results and provide exhaustive coverage across US, focusing on commute times
- Decomposition of difference by demographics
  - **Contribution:** Study new outcome and extend decomposition methods to study differences across locations
- Spatial stratification and urban form
  - **Contribution:** Urban form and housing prices are correlated with and drive our measure of spatial stratification

Literature & Contributions

- Document and quantify racialized difference in US commuting
  - **Contribution:** Update results and provide exhaustive coverage across US, focusing on commute times
- Decomposition of difference by demographics
  - **Contribution:** Study new outcome and extend decomposition methods to study differences across locations
- Spatial stratification and urban form
  - **Contribution:** Urban form and housing prices are correlated with and drive our measure of spatial stratification
- Employment suburbanization and Black suburbanization
  - Complement an exciting and growing literature (Aliprantis, Carroll, & Young 2019; Bartik & Mast 2021; Blair 2017; Miller 2022; Wiese 2005)

1. Methodology
2. Data
3. Aggregate Trends
4. Decomposition: ‘Individual’ Drivers of Racialized Difference
5. City-Level Drivers of Racialized Difference

Methodology – Aggregate Trends and Definitions

Evaluate average levels and changes in (i) **commute time** and (ii) **mode share** by **race**

Race: Focus on differential outcomes between Black and White commuters

- Black – identify as "Black" either alone or in combination with another race
- White – identify as "White" only
- Results similar if we instead compare Black and Non-Black commuters

Commute Time: Usual home→work travel time in minutes

Mode Share: Primary mode of transit (used most days/most distance)

- Automobile includes (motorcycle, taxi, and carpool)
- Bus includes (streetcar, trolleybus); Subway includes elevated; railroad is commuter rail
- Also: Bicycle; Walked only; and Other

Methodology – Regression Analysis

Explain *i*'s travel time, in commuting zone (CZ) *c* and year *t*:

$$\ln(\tau_{ict}) = \beta_t \mathbf{1}[\text{Black}_{ict}] + u_{ict} \quad (1)$$

$$\ln(\tau_{ict}) = \beta_t^* \mathbf{1}[\text{Black}_{ict}] + x_{ict}'\mu_t + \lambda_{ct} + u_{ict} \quad (2)$$

Include/exclude:

- CZ-fixed effects
  - Compares within CZ
- Covariates
  - Demographics/Educ, Trans. Mode, Job/Income

β_t, β_t^*: time-varying log difference

Cluster by CZ throughout paper

Notes & Caveats:

Interpretation: "controlling" for *x*?

- Discrimination or structural racism could drive different values of *x*
- Interpret as potential mechanisms

Selection:

- People are selecting LFP, mode, etc.
- λ_{ct} helps limit this...
- $\hat{\beta}_t$, $\hat{\beta}_t^*$ likely understate difference

Methodology – Decomposition

Decomposition framework

$$\ln(\tau_{ict}) = \alpha_t^W + x'_{ict}\mu_t^W + \lambda_{ct} + \epsilon_{ict}^W \qquad \text{if } \mathbf{1}[\text{Black}_{ict}] = 0$$
$$\ln(\tau_{ict}) = \alpha_t^B + x'_{ict}\mu_t^B + \lambda_{ct} + \epsilon_{ict}^B \qquad \text{if } \mathbf{1}[\text{Black}_{ict}] = 1$$

Regression-compatibility (Fortin 2008)

$$\mu^W = \mu^B = \mu \quad \Rightarrow \quad \alpha_t^B - \alpha_t^W = \Delta^{\text{Unexplained}} \quad = \quad \beta^*$$

Thus we have a Kitigawa-Oaxaca-Blinder decomposition:

$$\beta \quad = \quad \beta^* \quad + \quad \Delta^{\text{Explained}}$$
$$\Delta^{\text{Total}} \; = \; \Delta^{\text{Unexplained}} + \left(\Delta^{\text{Demog/Ed}} + \Delta^{\text{Transpo}} + \Delta^{\text{Work/Income}} + \Delta^{\text{Com. Zone}}\right)$$

Methodology – City-level Variation

Is there systematic, *city-level* variation in commuting difference?

Step 1: Estimate $\hat{\beta}_{ct}$: the **residual racialized difference** (RRD) in c

$$\ln(\tau_{ict}) = \beta_{ct}\mathbf{1}[\text{Black}_{ict}] + x'_{ict}\mu_{ct} + \lambda_{ct} + u_{ict}$$

- Similar to Eq. (1) except β & μ vary by c
- We show that both sources are components of $\Delta^{\text{Unexplained}}$

Methodology – City-level Variation

Is there systematic, *city-level* variation in commuting difference?

Step 1: Estimate $\hat{\beta}_{ct}$: the **residual racialized difference** (RRD) in c

$$\ln(\tau_{ict}) = \beta_{ct}\mathbf{1}[\text{Black}_{ict}] + x'_{ict}\mu_{ct} + \lambda_{ct} + u_{ict}$$

- Similar to Eq. (1) except β & μ vary by c
- We show that both sources are components of $\Delta^{\text{Unexplained}}$

Step 2: Estimate city-level correlates of RRD (e.g., urban form, segregation)

$$\hat{\beta}_{ct} = z'_{ct}\gamma + D_c + T_t + e_{ct}$$

- Specifications with and without fixed effects

1. Methodology
2. Data
3. Aggregate Trends
4. Decomposition: ‘Individual’ Drivers of Racialized Difference
5. City-Level Drivers of Racialized Difference

Data – Primary Source

Census/ACS, 1980–2019; sample consists of all *commuters*

- Journey to Work questions ask about race and commute time/mode
- We assign to consistent commuting zones (CZs) (Autor & Dorn '13)
  - Lightly modify to bring together large markets, e.g., DFW, NYC/Newark
- Often focus on year bins: 1980, 1990, 2000, 2005–11, 2012–19
- Before 2000, race in the Census was *univariate*
  - In 2000 and later, race could be multi-dimensional
  - Selection of multiple races increase substantially in 2010s

Data – Primary Source

Census/ACS, 1980–2019; sample consists of all *commuters*

- Journey to Work questions ask about race and commute time/mode
- We assign to consistent commuting zones (CZs) (Autor & Dorn '13)
  - Lightly modify to bring together large markets, e.g., DFW, NYC/Newark
- Often focus on year bins: 1980, 1990, 2000, 2005–11, 2012–19
- Before 2000, race in the Census was *univariate*
  - In 2000 and later, race could be multi-dimensional
  - Selection of multiple races increase substantially in 2010s
- Extend back to 1960 for aggregate mode share
- Some specs. include residential PUMA geographies starting in 2000

Data – Covariates

Observable covariates and groups of covariates in Census/ACS (harmonized)

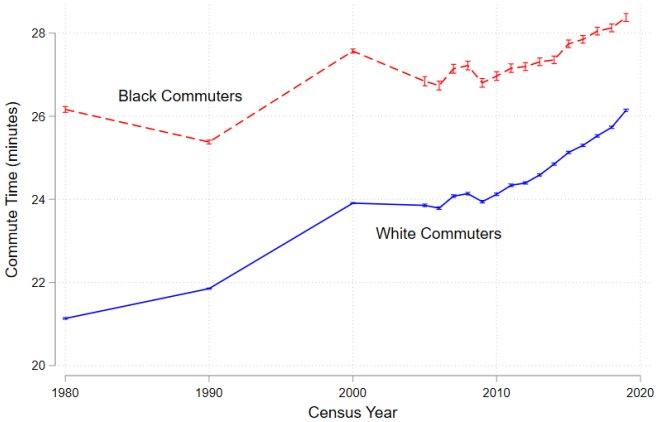
- *Commuting Zone*: fixed effects for CZ
- *Demographics & Education*:
  - sex
  - indicators for education (less than high school, high school, college graduate, and masters or higher)
  - a quadratic in age
  - marital status
  - head of household
  - indicators for numbers of children (zero, one or two, and three or more)
- *Transportation Mode* indicators:
  - private motor vehicle (including motorcycle, taxi, and carpool)
  - bus or streetcar
  - subway or elevated
  - railroad (commuter rail)
  - bicycle; walked only; and other
- *Work & Income*:
  - indicator for zero income
  - log income (set to 0 if zero income)
  - indicators for industry
  - indicators for occupation

Data – Secondary Sources

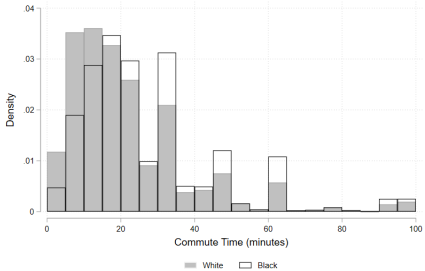
- NHGIS for finer (census tract/ZCTA) geographic aggregates
  - Geonormalize to study average tract-level commuting time (+ tract FEs)
  - Use to create city-specific measures of urban form (segregation, centrality)
  - ... but not microdata
- Zip Code Business Patterns for spatial dist. of work locations
  - Colocation of jobs and housing, employment concentration
- Miles of highway (Baum-Snow 2007)

1. Methodology
2. Data
3. Aggregate Trends
4. Decomposition: ‘Individual’ Drivers of Racialized Difference
5. City-Level Drivers of Racialized Difference

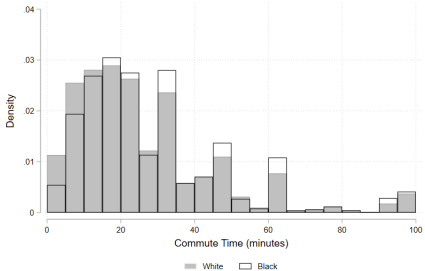
Aggregate Differences in Commute Time



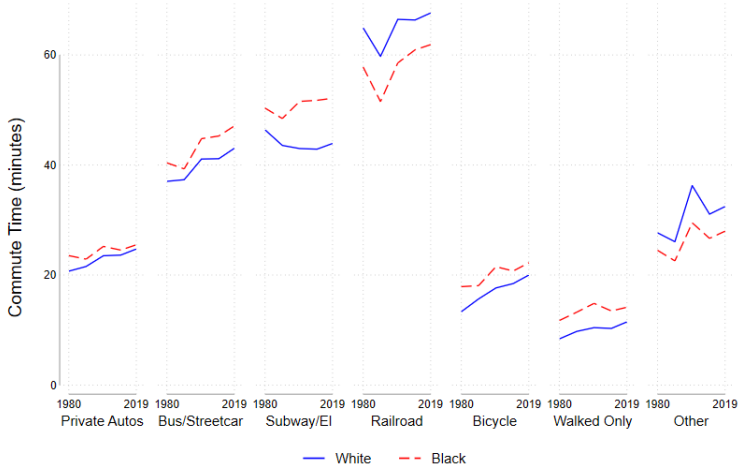
1980



2012–19



Aggregate Differences in Commute Time by Mode



Black commuters face longer commutes for most modes

- Especially transit

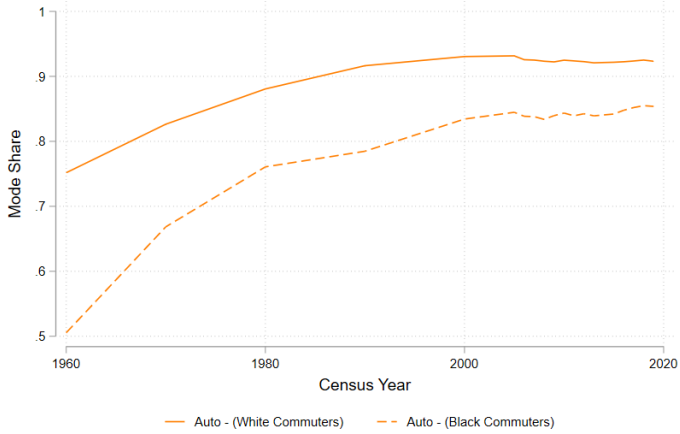
Mode is an important channel

Railroad/Other are counter-examples (but marginal)

Aggregate Differences in Mode Share

Large increase in auto commutes, 1960–2019

- Primarily at the expense of Bus/Streetcar use by Black commuters Transit Share
- Also substantial reduction of Walking for all commuters Walk/Other Share



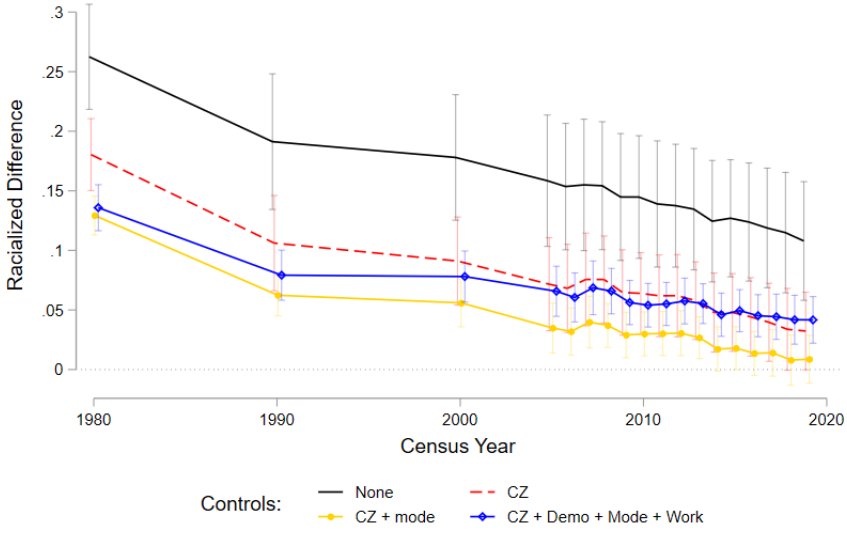
1. Methodology
2. Data
3. Aggregate Trends
4. Decomposition: ‘Individual’ Drivers of Racialized Difference
5. City-Level Drivers of Racialized Difference

Baseline Results (1980 & 2012–19)

	$\ln(\tau_{ict})$					
	(1)	(2)	(3)	(4)	(5)	(6)
$1[\text{Black}] \times t_{1980}$	0.263***	0.180***	0.198***	0.129***	0.139***	0.136***
	(0.022)	(0.015)	(0.016)	(0.008)	(0.008)	(0.010)
$1[\text{Black}] \times t_{2012-19}$	0.124***	0.046***	0.070***	0.018*	0.037***	0.049***
	(0.025)	(0.017)	(0.017)	(0.009)	(0.009)	(0.009)
Year Bin × CZ FEs	-	Y	Y	Y	Y	Y
Controls – Demog. & Edu.	-	-	Y	-	Y	Y
Controls – Trans. Mode	-	-	-	Y	Y	Y
Controls – Work & Income	-	-	-	-	-	Y

- Black commutes 30% longer in 1980, 13% longer in 2012–19 (unconditional)
- CZ fixed effects reduce this by about 8 log points
- Transportation mode seems to be explanatory as well

Baseline Results



Decomposition

Non-sequential decomposition (Gelbach '16)

	Δ_t^{Total}	$\Delta_t^{\text{Unexpl.}}$	$\Delta_t^{\text{Explained}}$			
			Δ_t^{Demog}	$\Delta_t^{\text{Tr. Mode}}$	$\Delta_t^{\text{Work/Inc}}$	Δ_t^{CZ}
Decomposition						
$1[\text{Black}] \times t_{1980}$	0.263	0.136	-0.008	0.073	-0.001	0.062
		51.7%	-3.0%	27.8%	-0.2%	23.7%
$1[\text{Black}] \times t_{1990}$	0.191	0.079	-0.009	0.063	-0.007	0.065
		41.4%	-5.0%	32.9%	-3.4%	34.0%
$1[\text{Black}] \times t_{2000}$	0.178	0.078	-0.008	0.050	-0.011	0.069
		43.9%	-4.6%	28.1%	-6.3%	39.0%
$1[\text{Black}] \times t_{2005\text{—}11}$	0.150	0.061	-0.009	0.049	-0.014	0.063
		40.5%	-6.1%	33.0%	-9.5%	42.1%
$1[\text{Black}] \times t_{2012\text{—}19}$	0.124	0.049	-0.008	0.040	-0.019	0.063
		39.2%	-6.6%	32.5%	-15.4%	50.4%
Components of Change						
$\frac{\Delta^k_{1980} - \Delta^k_{2012\text{—}19}}{\Delta_{1980} - \Delta_{2012\text{—}19}}$	-	62.6%	0.0%	23.7%	12.9%	-0.7%

Decomposition

Non-sequential decomposition (Gelbach '16)

	Δ_t^{Total}	$\Delta_t^{\text{Unexpl.}}$	$\Delta_t^{\text{Explained}}$			
			Δ_t^{Demog}	$\Delta_t^{\text{Tr. Mode}}$	$\Delta_t^{\text{Work/Inc}}$	Δ_t^{CZ}
Decomposition						
$1[\text{Black}] \times t_{1980}$	0.263	0.136	-0.008	0.073	-0.001	0.062
		51.7%	-3.0%	27.8%	-0.2%	23.7%
$1[\text{Black}] \times t_{1990}$	0.191	0.079	-0.009	0.063	-0.007	0.065
		41.4%	-5.0%	32.9%	-3.4%	34.0%
$1[\text{Black}] \times t_{2000}$	0.178	0.078	-0.008	0.050	-0.011	0.069
		43.9%	-4.6%	28.1%	-6.3%	39.0%
$1[\text{Black}] \times t_{2005-11}$	0.150	0.061	-0.009	0.049	-0.014	0.063
		40.5%	-6.1%	33.0%	-9.5%	42.1%
$1[\text{Black}] \times t_{2012-19}$	0.124	0.049	-0.008	0.040	-0.019	0.063
		39.2%	-6.6%	32.5%	-15.4%	50.4%
Components of Change						
$\frac{\Delta^k_{1980} - \Delta^k_{2012-19}}{\Delta_{1980} - \Delta_{2012-19}}$	-	62.6%	0.0%	23.7%	12.9%	-0.7%

- City plays constant role in level

Decomposition

Non-sequential decomposition (Gelbach '16)

	Δ_t^{Total}	$\Delta_t^{\text{Unexpl.}}$	$\Delta_t^{\text{Explained}}$			
			Δ_t^{Demog}	$\Delta_t^{\text{Tr. Mode}}$	$\Delta_t^{\text{Work/Inc}}$	Δ_t^{CZ}
Decomposition						
$1[\text{Black}] \times t_{1980}$	0.263	0.136	-0.008	0.073	-0.001	0.062
		51.7%	-3.0%	27.8%	-0.2%	23.7%
$1[\text{Black}] \times t_{1990}$	0.191	0.079	-0.009	0.063	-0.007	0.065
		41.4%	-5.0%	32.9%	-3.4%	34.0%
$1[\text{Black}] \times t_{2000}$	0.178	0.078	-0.008	0.050	-0.011	0.069
		43.9%	-4.6%	28.1%	-6.3%	39.0%
$1[\text{Black}] \times t_{2005-11}$	0.150	0.061	-0.009	0.049	-0.014	0.063
		40.5%	-6.1%	33.0%	-9.5%	42.1%
$1[\text{Black}] \times t_{2012-19}$	0.124	0.049	-0.008	0.040	-0.019	0.063
		39.2%	-6.6%	32.5%	-15.4%	50.4%
Components of Change						
$\frac{\Delta^k_{1980} - \Delta^k_{2012-19}}{\Delta_{1980} - \Delta_{2012-19}}$	-	62.6%	0.0%	23.7%	12.9%	-0.7%

- City plays constant role in level
- Mode plays constant relative role

Decomposition

Non-sequential decomposition (Gelbach '16)

	Δ_t^{Total}	$\Delta_t^{\text{Unexpl.}}$	$\Delta_t^{\text{Explained}}$			
			Δ_t^{Demog}	$\Delta_t^{\text{Tr. Mode}}$	$\Delta_t^{\text{Work/Inc}}$	Δ_t^{CZ}
Decomposition						
$1[\text{Black}] \times t_{1980}$	0.263	0.136	-0.008	0.073	-0.001	0.062
		51.7%	-3.0%	27.8%	-0.2%	23.7%
$1[\text{Black}] \times t_{1990}$	0.191	0.079	-0.009	0.063	-0.007	0.065
		41.4%	-5.0%	32.9%	-3.4%	34.0%
$1[\text{Black}] \times t_{2000}$	0.178	0.078	-0.008	0.050	-0.011	0.069
		43.9%	-4.6%	28.1%	-6.3%	39.0%
$1[\text{Black}] \times t_{2005-11}$	0.150	0.061	-0.009	0.049	-0.014	0.063
		40.5%	-6.1%	33.0%	-9.5%	42.1%
$1[\text{Black}] \times t_{2012-19}$	0.124	0.049	-0.008	0.040	-0.019	0.063
		39.2%	-6.6%	32.5%	-15.4%	50.4%
Components of Change						
$\frac{\Delta^k_{1980} - \Delta^k_{2012-19}}{\Delta_{1980} - \Delta_{2012-19}}$	-	62.6%	0.0%	23.7%	12.9%	-0.7%

- City plays constant role in level
- Mode plays constant relative role
- Negative selection on work and demographics
  - $\rho(\ln w, \ln \tau) > 0$...

Decomposition

Non-sequential decomposition (Gelbach '16)

	Δ_t^{Total}	$\Delta_t^{\text{Unexpl.}}$	$\Delta_t^{\text{Explained}}$			
			Δ_t^{Demog}	$\Delta_t^{\text{Tr. Mode}}$	$\Delta_t^{\text{Work/Inc}}$	Δ_t^{CZ}
Decomposition						
$1[\text{Black}] \times t_{1980}$	0.263	0.136 51.7%	-0.008 -3.0%	0.073 27.8%	-0.001 -0.2%	0.062 23.7%
$1[\text{Black}] \times t_{1990}$	0.191	0.079 41.4%	-0.009 -5.0%	0.063 32.9%	-0.007 -3.4%	0.065 34.0%
$1[\text{Black}] \times t_{2000}$	0.178	0.078 43.9%	-0.008 -4.6%	0.050 28.1%	-0.011 -6.3%	0.069 39.0%
$1[\text{Black}] \times t_{2005-11}$	0.150	0.061 40.5%	-0.009 -6.1%	0.049 33.0%	-0.014 -9.5%	0.063 42.1%
$1[\text{Black}] \times t_{2012-19}$	0.124	0.049 39.2%	-0.008 -6.6%	0.040 32.5%	-0.019 -15.4%	0.063 50.4%
Components of Change						
$\frac{\Delta^k_{1980} - \Delta^k_{2012-19}}{\Delta_{1980} - \Delta_{2012-19}}$	-	62.6%	0.0%	23.7%	12.9%	-0.7%

- City plays constant role in level
- Mode plays constant relative role
- Negative selection on work and demographics
  - $\rho(\ln w, \ln \tau) > 0$...
- Large portion unexplained

Decomposition

Non-sequential decomposition (Gelbach '16)

	Δ_t^{Total}	$\Delta_t^{\text{Unexpl.}}$	$\Delta_t^{\text{Explained}}$			
			Δ_t^{Demog}	$\Delta_t^{\text{Tr. Mode}}$	$\Delta_t^{\text{Work/Inc}}$	Δ_t^{CZ}
Decomposition						
$1[\text{Black}] \times t_{1980}$	0.263	0.136 51.7%	-0.008 -3.0%	0.073 27.8%	-0.001 -0.2%	0.062 23.7%
$1[\text{Black}] \times t_{1990}$	0.191	0.079 41.4%	-0.009 -5.0%	0.063 32.9%	-0.007 -3.4%	0.065 34.0%
$1[\text{Black}] \times t_{2000}$	0.178	0.078 43.9%	-0.008 -4.6%	0.050 28.1%	-0.011 -6.3%	0.069 39.0%
$1[\text{Black}] \times t_{2005-11}$	0.150	0.061 40.5%	-0.009 -6.1%	0.049 33.0%	-0.014 -9.5%	0.063 42.1%
$1[\text{Black}] \times t_{2012-19}$	0.124	0.049 39.2%	-0.008 -6.6%	0.040 32.5%	-0.019 -15.4%	0.063 50.4%
Components of Change						
$\frac{\Delta^k_{1980} - \Delta^k_{2012-19}}{\Delta_{1980} - \Delta_{2012-19}}$	-	62.6%	0.0%	23.7%	12.9%	-0.7%

- City plays constant role in level
- Mode plays constant relative role
- Negative selection on work and demographics
  - $\rho(\ln w, \ln \tau) > 0$...
- Large portion unexplained
- Partial Convergence due to Mode and Unexplained

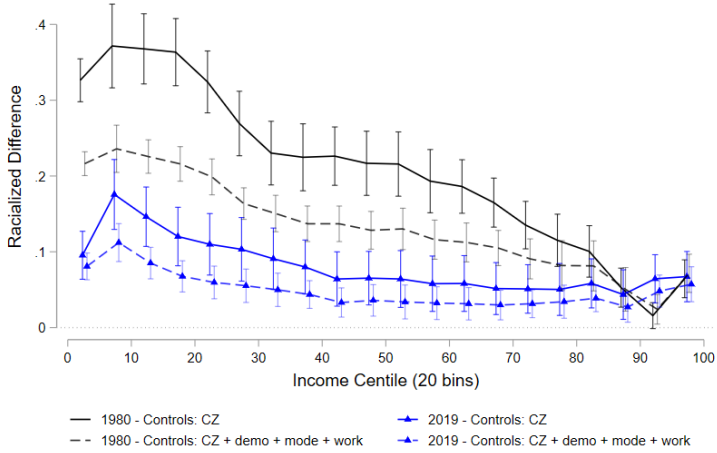
Extensions

Test heterogeneity & linearity:

1. Heterogeneity by income
2. Differences by mode

Does residential neighborhood explain the difference?

3. Include PUMAs as controls in 2000 and later
4. Census-tract-level outcomes & FEs (but no microdata)



Extensions

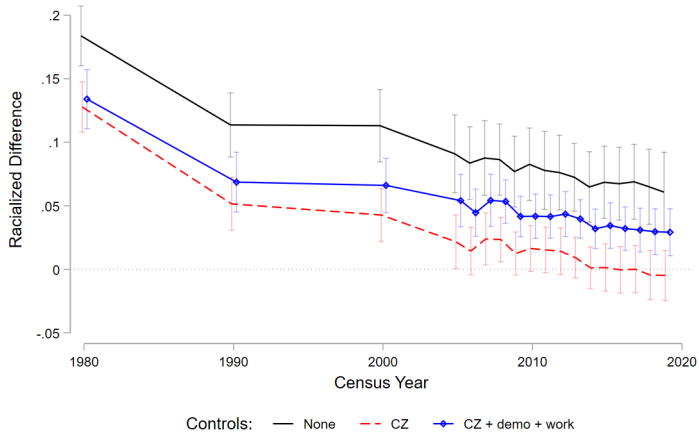
Car:

Test heterogeneity & linearity:

1. Heterogeneity by income
2. Differences by mode

Does residential neighborhood explain the difference?

3. Include PUMAs as controls in 2000 and later
4. Census-tract-level outcomes & FEs (but no microdata)



Extensions

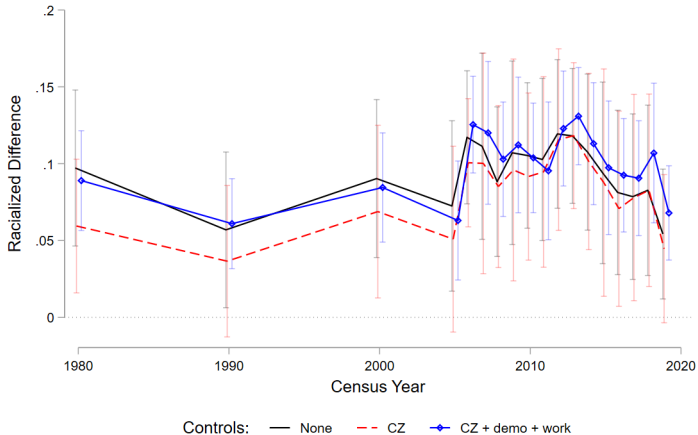
Test heterogeneity & linearity:

1. Heterogeneity by income
2. Differences by mode

Does residential neighborhood explain the difference?

3. Include PUMAs as controls in 2000 and later
4. Census-tract-level outcomes & FEs (but no microdata)

Bus:



Extensions

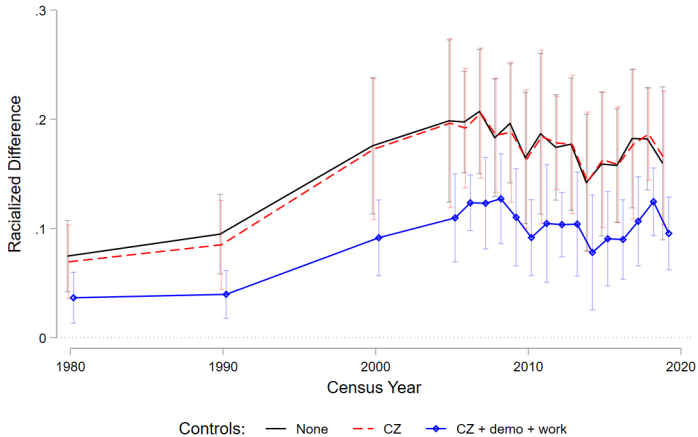
Test heterogeneity & linearity:

1. Heterogeneity by income
2. Differences by mode

Does residential neighborhood explain the difference?

3. Include PUMAs as controls in 2000 and later
4. Census-tract-level outcomes & FEs (but no microdata)

Subway:



Extensions

Test heterogeneity & linearity:

1. Heterogeneity by income
2. Differences by mode

Does residential neighborhood explain the difference?

3. Include PUMAs as controls in 2000 and later
4. Census-tract-level outcomes & FEs (but no microdata)

	All Modes (1)	Car (2)	Bus (3)	Subway (4)	Walk (5)
A. No PUMAs					
$1[\text{Black}] \times t_{2000}$	0.078***	0.066***	0.085***	0.091***	0.291***
	(0.011)	(0.011)	(0.018)	(0.018)	(0.022)
$1[\text{Black}] \times t_{2005-11}$	0.061***	0.047***	0.102***	0.114***	0.208***
	(0.010)	(0.009)	(0.016)	(0.019)	(0.023)
$1[\text{Black}] \times t_{2012-19}$	0.049***	0.035***	0.104***	0.102***	0.172***
	(0.009)	(0.008)	(0.016)	(0.019)	(0.018)
B. With PUMA-Year FEs					
$1[\text{Black}] \times t_{2000}$	0.076***	0.069***	0.069***	0.022***	0.255***
	(0.006)	(0.006)	(0.012)	(0.007)	(0.016)
$1[\text{Black}] \times t_{2005-11}$	0.060***	0.053***	0.079***	0.036***	0.196***
	(0.005)	(0.006)	(0.007)	(0.010)	(0.013)
$1[\text{Black}] \times t_{2012-19}$	0.043***	0.034***	0.071***	0.033***	0.153***
	(0.004)	(0.004)	(0.008)	(0.009)	(0.012)
N	37 mil.	35 mil.	527k	303k	1 mil.

Extensions

Test heterogeneity & linearity:

1. Heterogeneity by income
2. Differences by mode

Does residential neighborhood explain the difference?

3. Include PUMAs as controls in 2000 and later
4. Census-tract-level outcomes & FEs (but no microdata)

$$\ln(\bar{\tau}_{act}) = \beta_t^* s_{act}^{\text{Black}} + \bar{x}'_{act}\mu + \xi_a + \lambda_{ct} + u_{act}$$

	(1)	(2)	(3)
$s_{act}^{\text{Black}} \times t_{1980}$	0.245***	0.129***	0.063***
	(0.042)	(0.024)	(0.016)
$s_{act}^{\text{Black}} \times t_{1990}$	0.179***	0.040	0.021
	(0.046)	(0.031)	(0.014)
$s_{act}^{\text{Black}} \times t_{2000}$	0.197***	0.073*	0.086***
	(0.047)	(0.035)	(0.012)
$s_{act}^{\text{Black}} \times t_{2006\text{—}10}$	0.132**	0.014	0.043***
	(0.047)	(0.035)	(0.011)
$s_{act}^{\text{Black}} \times t_{2014\text{—}18}$	0.112*	-0.004	0.044***
	(0.049)	(0.038)	(0.012)
N	346,631	346,522	346,478
Year Bin × CZ FEs	-	Y	Y
Share Transit in Tract	-	-	Y
Tract FEs	-	-	Y

$\bar{\tau}$ is agg. minutes/commuters, or prediction if only binned times are available

Summary of 'Individual' Results

- Large gap in average commute time by race
  - Racialized difference mostly reflects city, mode, and unexplained factors
- Partial convergence between 1980 and 2019
  - Largely explained by mode (partial convergence to automobile use)
  - Substantial portion of difference unexplained today (41%)
- Much (63%) of this partial convergence is due to unobserved factors
- Racialized difference is present
  - Across the income spectrum
  - For users of all modes, though less so for automobile commuters
  - Even conditional on PUMA/neighborhood fixed effects

1. Methodology
2. Data
3. Aggregate Trends
4. Decomposition: ‘Individual’ Drivers of Racialized Difference
5. City-Level Drivers of Racialized Difference

City-level Heterogeneity

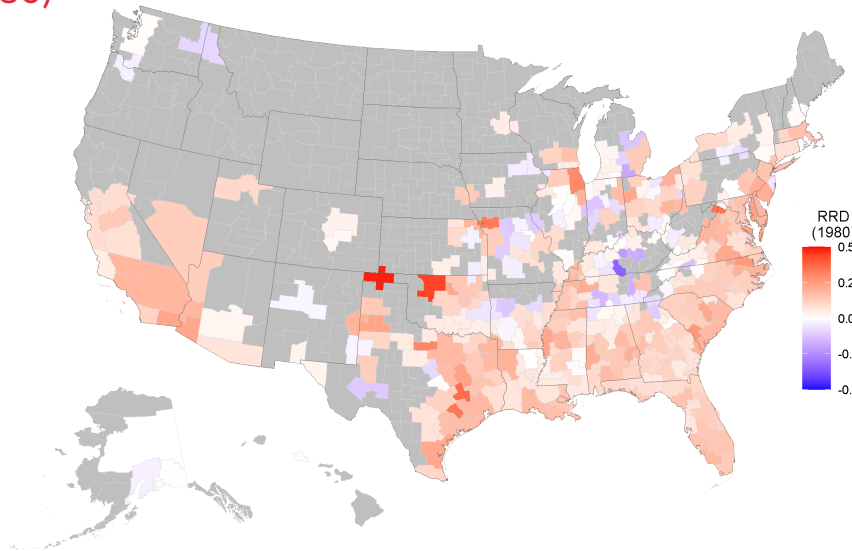
What correlates with (or drives) *city-level variation* in this difference? Two-step approach:

$$\ln(\tau_{ict}) = \beta_{ct}\mathbf{1}[\text{Black}_{ict}] + x'_{ict}\mu_{ct} + \lambda_{ct} + u_{ict}$$
$$\hat{\beta}_{ct} = z'_{ct}\gamma + D_c + T_t + e_{ct}$$

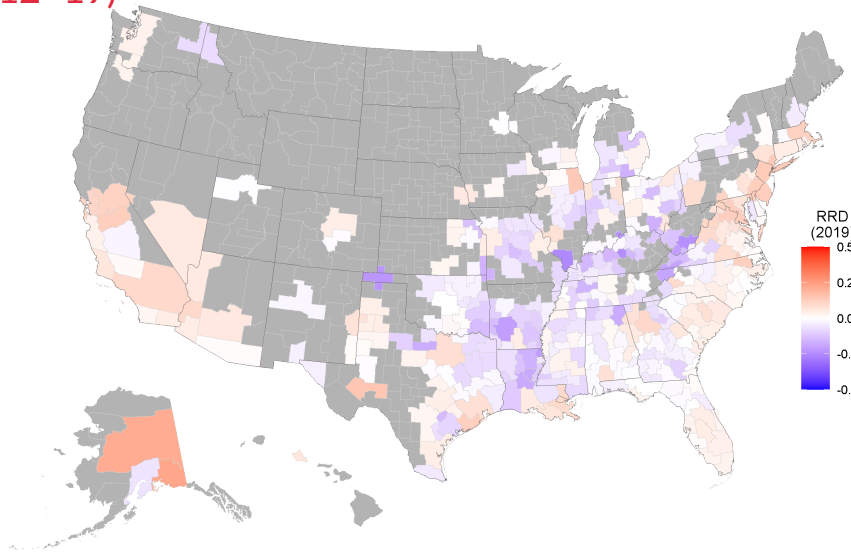
$\hat{\beta}_{ct}$ is the **residual racialized difference** (RRD) in commute time

- RRD contributes to $\Delta^{\text{Unexplained}}_{\{t\}}$
- Dealing with generated β and heteroskedasticity. Drop
  - i) CZs with $<$1k commuters,
  - ii) CZs with $<$50 unique Black commuter Census respondents.
  - Weight second stage by number of Black commuters
- Cluster SEs by CZ

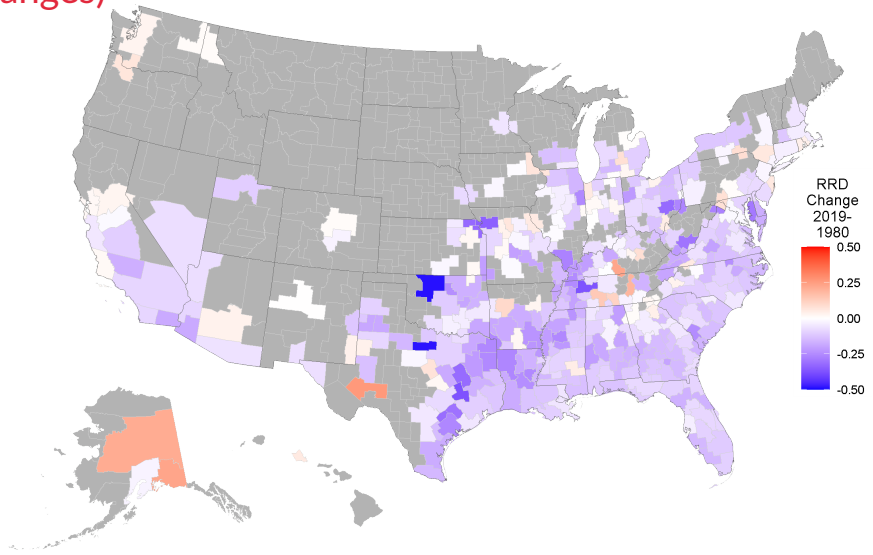
Map (1980)



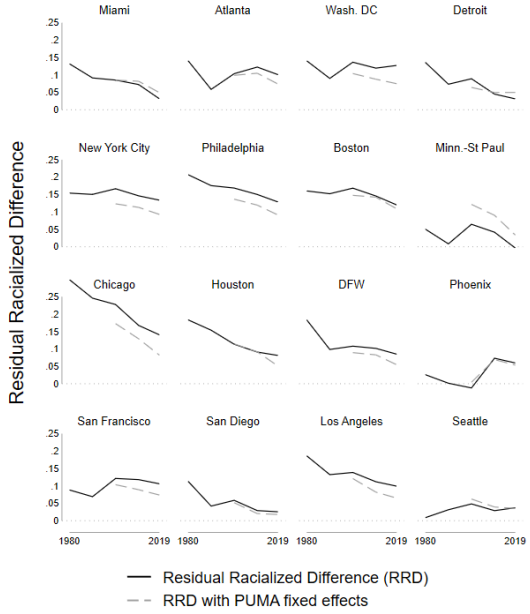
Map (2012–19)



Map (Changes)



RRD in 16 Cities



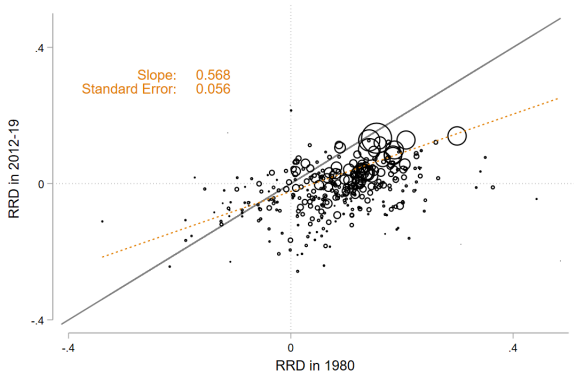
Summary and Persistence

Summary of 341 consistent RRDs

	Mean	SD	Min	Max
1980	0.131	0.072	-0.339	0.485
1990	0.070	0.072	-0.326	0.246
2000	0.068	0.077	-0.412	0.247
2005-11	0.053	0.073	-0.384	0.220
2012-19	0.032	0.070	-0.257	0.230

Contribution to unexplained difference:

	$\Delta^{\text{Unexplained}}$	$\Delta^{\text{Unexplained, het. CZ}}$
1980	0.136	0.105
2012-19	0.049	0.038

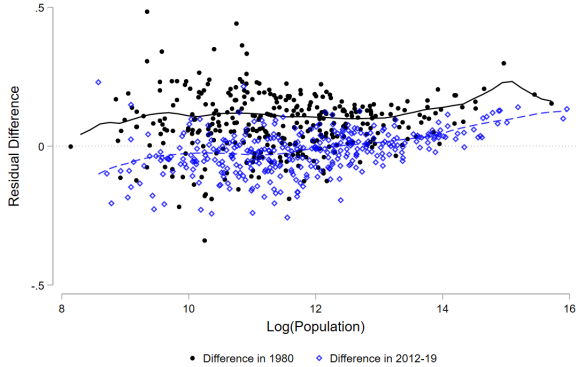


Declining mean, but not much decline in SD

- Relatively high but not uniform persistence over 40-year interval

Correlates - City Size

- Population increasingly predicts a larger RRD
- Population growth somewhat predictive, but less so for big cities
- Black share of population predicting less and less



	1980		2012–19		Panel (+FEs)	
Ln(Pop)	0.025***	0.033**	0.037***	0.046***	0.040+	0.022
	(0.007)	(0.013)	(0.002)	(0.004)	(0.023)	(0.028)
Cities	All	>200k	All	>200k	All	>200k
N	341	90	341	90	1705	450
R^2	0.304	0.336	0.621	0.657	0.861	0.883

All models include Black share of commuting population as control.

Urban Form and RRD

Correlates of RRD

$$\hat{\beta}_{ct} = z'_{ct}\gamma + D_c + T_t + e_{ct}$$

Several measure of urban form

- Segregation, centrality, job/residence colocation, transportation
- (Un?)Conditional on population

Focus on big cities ($>$ 200k commuters, i.e., 90 largest CZs)

- Weak relationships in small cities
- E.g., Birmingham vs Chicago

Correlates - Urban Form (Larger Cities)

	TWFE correlation of … with RRD						
	Dissimi-larity (1)	Black Empl. Conc. (2)	White Empl. Conc. (3)	Cen-trality (4)	Log Hwy Miles (5)	Transit Share (6)	Ave. Car Time (7)
Panel A. No Controls							
Measure	0.2448*	0.2379**	-0.2927+	0.0098	-0.0791**	0.4587**	0.0056+
	(0.1160)	(0.0707)	(0.1692)	(0.0801)	(0.0285)	(0.1716)	(0.0032)
Panel B. Controlling for Log Population							
Measure	0.2863*	0.2282**	-0.2392	0.0404	-0.0710**	0.4604**	0.0047
	(0.1147)	(0.0731)	(0.1559)	(0.0696)	(0.0245)	(0.1570)	(0.0033)
N	450	360	360	450	264	450	450
Sample Years	’80-’19	’90-’19	’90-’19	’80-’19	’80-’00	’80-’19	’80-’19

- More segregated CZs (↑ dissimilarity) have higher RRD

Correlates - Urban Form (Larger Cities)

	TWFE correlation of … with RRD						
	Dissimi-larity (1)	Black Empl. Conc. (2)	White Empl. Conc. (3)	Cen-trality (4)	Log Hwy Miles (5)	Transit Share (6)	Ave. Car Time (7)
Panel A. No Controls							
Measure	0.2448*	0.2379**	-0.2927+	0.0098	-0.0791**	0.4587**	0.0056+
	(0.1160)	(0.0707)	(0.1692)	(0.0801)	(0.0285)	(0.1716)	(0.0032)
Panel B. Controlling for Log Population							
Measure	0.2863*	0.2282**	-0.2392	0.0404	-0.0710**	0.4604**	0.0047
	(0.1147)	(0.0731)	(0.1559)	(0.0696)	(0.0245)	(0.1570)	(0.0033)
N	450	360	360	450	264	450	450
Sample Years	'80-'19	'90-'19	'90-'19	'80-'19	'80-'00	'80-'19	'80-'19

- Colocation of jobs and Black residential location (↓ Black Empl. Conc.) reduces RRD
  - Empl. concentration is GINI(jobs, residential location by race) by zip code (Bento et al. '05)

Correlates - Urban Form (Larger Cities)

	TWFE correlation of … with RRD						
	Dissimi-larity (1)	Black Empl. Conc. (2)	White Empl. Conc. (3)	Cen-trality (4)	Log Hwy Miles (5)	Transit Share (6)	Ave. Car Time (7)
Panel A. No Controls							
Measure	0.2448*	0.2379**	-0.2927+	0.0098	-0.0791**	0.4587**	0.0056+
	(0.1160)	(0.0707)	(0.1692)	(0.0801)	(0.0285)	(0.1716)	(0.0032)
Panel B. Controlling for Log Population							
Measure	0.2863*	0.2282**	-0.2392	0.0404	-0.0710**	0.4604**	0.0047
	(0.1147)	(0.0731)	(0.1559)	(0.0696)	(0.0245)	(0.1570)	(0.0033)
N	450	360	360	450	264	450	450
Sample Years	'80-'19	'90-'19	'90-'19	'80-'19	'80-'00	'80-'19	'80-'19

- Colocation of jobs and White residential location (↓ White Empl. Conc.) increases RRD
  - Empl. concentration is GINI(jobs, residential location by race) by zip code (Bento et al. '05)

Correlates - Urban Form (Larger Cities)

	TWFE correlation of … with RRD						
	Dissimi-larity (1)	Black Empl. Conc. (2)	White Empl. Conc. (3)	Cen-trality (4)	Log Hwy Miles (5)	Transit Share (6)	Ave. Car Time (7)
Panel A. No Controls							
Measure	0.2448*	0.2379**	-0.2927+	0.0098	-0.0791**	0.4587**	0.0056+
	(0.1160)	(0.0707)	(0.1692)	(0.0801)	(0.0285)	(0.1716)	(0.0032)
Panel B. Controlling for Log Population							
Measure	0.2863*	0.2282**	-0.2392	0.0404	-0.0710**	0.4604**	0.0047
	(0.1147)	(0.0731)	(0.1559)	(0.0696)	(0.0245)	(0.1570)	(0.0033)
N	450	360	360	450	264	450	450
Sample Years	’80-’19	’90-’19	’90-’19	’80-’19	’80-’00	’80-’19	’80-’19

- Centrality does not seem to play a large role

Correlates - Urban Form (Larger Cities)

	TWFE correlation of … with RRD						
	Dissimi-larity (1)	Black Empl. Conc. (2)	White Empl. Conc. (3)	Cen-trality (4)	Log Hwy Miles (5)	Transit Share (6)	Ave. Car Time (7)
Panel A. No Controls							
Measure	0.2448*	0.2379**	-0.2927+	0.0098	-0.0791**	0.4587**	0.0056+
	(0.1160)	(0.0707)	(0.1692)	(0.0801)	(0.0285)	(0.1716)	(0.0032)
Panel B. Controlling for Log Population							
Measure	0.2863*	0.2282**	-0.2392	0.0404	-0.0710**	0.4604**	0.0047
	(0.1147)	(0.0731)	(0.1559)	(0.0696)	(0.0245)	(0.1570)	(0.0033)
N	450	360	360	450	264	450	450
Sample Years	’80-’19	’90-’19	’90-’19	’80-’19	’80-’00	’80-’19	’80-’19

- “Faster” cities (more highway, less transit, faster car) have smaller RRD

Ingredients of Stratification?

Bigger and slower cities see higher RRDs

- In smaller (or faster) places, job access is more equidistant in time
- Increasingly, *only* big cities systematically see higher RRD

Declining segregation $\leftrightarrow$ decreasing RRD

Employment concentration

From map: Coastal cities see persistent RRD

Stratification now occurs via housing prices $\Rightarrow$ investigate role of housing prices

How Housing Prices Might Impact Stratification?

1. Housing price dispersion ↑ since 1970s due to worker sorting (Van Nieuwerburgh & Weill '10)
2. Housing demand ⇒ spatial neighborhood change (low-income n'hoods near high-income n'hoods shift to high income) (Guerrieri, Hartley, Hurst '13)
   - Access is a persistent 'second-nature' neighborhood amenity (e.g., Cronon '91)
   - Big, expensive cities features lots of variation in job access
3. High prices crowd out low-income households from 'superstar' areas within MSAs (Gyourko, Mayer, Sinai '13)

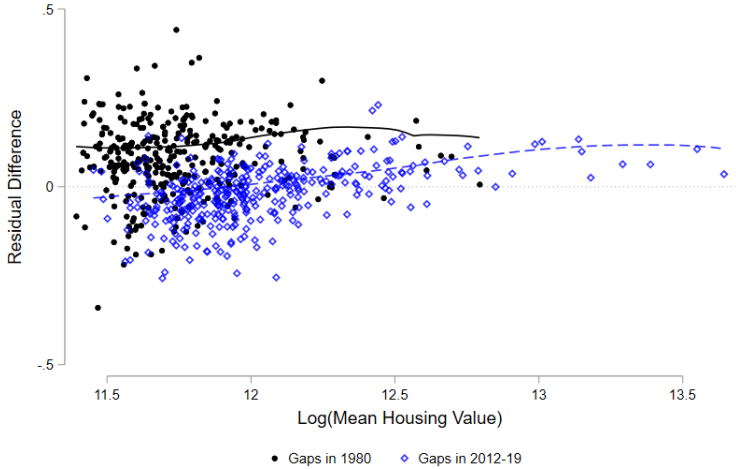
How Housing Prices Might Impact Stratification?

1. Housing price dispersion ↑ since 1970s due to worker sorting (Van Nieuwerburgh & Weill '10)
2. Housing demand ⇒ spatial neighborhood change (low-income n'hoods near high-income n'hoods shift to high income) (Guerrieri, Hartley, Hurst '13)
   - Access is a persistent 'second-nature' neighborhood amenity (e.g., Cronon '91)
   - Big, expensive cities features lots of variation in job access
3. High prices crowd out low-income households from 'superstar' areas within MSAs (Gyourko, Mayer, Sinai '13)

Meanwhile...

- Evolving preferences for jobs and time use (Edlund, Machado, & Sviatschi '21; Su '21)
- Geography lurking in the background (Saiz '10; Lee & Lin '18; Saiz & Wang '21)
- Inelastic supply likely exacerbates access issues even while prices increase
- Substantial accumulated wealth differences by race (Kuhn, Schularick, & Steins' 20)
- Steering & discrimination in housing markets (Christensen & Timmins '21)

Housing Prices and RRD



Details of Housing IV

Look for relationship between Δ housing prices and RRD

- Concerns about reverse causality and confounding factors (e.g., land use regs, prod. shocks to clusters)
- Turn to IV that exploits varied exposure to regional housing cycles (Guren et al. '22)

$$P_{cdt} = \delta_c \bar{P}_{(-c)dt} + \psi_0 \hat{\beta}_{ct} + \psi_1 m_{cdt} + \phi_c t + D_c + \epsilon_{cdt} \qquad \text{(Step 1)}$$

- P_{cdt} is log mean housing price in CZ c in Census division d in year-bin t
- $\bar{P}_{(-c)dt}$ is the leave-c-out log mean housing price in the Census division
- $\hat{\delta}_c \bar{P}_{(-c)dt}$ measures local response to reg. price movements $\rightarrow$ time-varying IV

Details of Housing IV

Look for relationship between Δ housing prices and RRD

- Concerns about reverse causality and confounding factors (e.g., land use regs, prod. shocks to clusters)
- Turn to IV that exploits varied exposure to regional housing cycles (Guren et al. '22)

$$P_{cdt} = \delta_c \bar{P}_{(-c)dt} + \psi_0 \hat{\beta}_{ct} + \psi_1 m_{cdt} + \phi_c t + D_c + \epsilon_{cdt} \quad \text{(Step 1)}$$

$$\hat{\beta}_{ct} = \gamma P_{cdt} + D_c + T_t + e_{ct} \quad \text{(Step 2)}$$

$$\mathbb{E}[\hat{\delta}_c \bar{P}_{(-c)dt} \times e_{ct}] = 0$$

- P_{cdt} is log mean housing price in CZ c in Census division d in year-bin t
- $\bar{P}_{(-c)dt}$ is the leave-c-out log mean housing price in the Census division
- $\hat{\delta}_c \bar{P}_{(-c)dt}$ measures local response to reg. price movements $\rightarrow$ time-varying IV

Details of Housing IV

Look for relationship between Δ housing prices and RRD

- Concerns about reverse causality and confounding factors (e.g., land use regs, prod. shocks to clusters)
- Turn to IV that exploits varied exposure to regional housing cycles (Guren et al. '22)

$$P_{cdt} = \delta_c \bar{P}_{(-c)dt} + \psi_0 \hat{\beta}_{ct} + \psi_1 m_{cdt} + \phi_c t + D_c + \epsilon_{cdt} \quad \text{(Step 1)}$$

$$\hat{\beta}_{ct} = \gamma P_{cdt} + D_c + T_t + e_{ct} \quad \text{(Step 2)}$$

$$\mathbb{E}[\hat{\delta}_c \bar{P}_{(-c)dt} \times e_{ct}] = 0$$

Identification requires there be no unobserved factor that:

i. is correlated with regional house price movements, and
ii. systematically has greater/lesser impact CZs more sensitive to regional demand shocks

- conditional on CZ-specific trends

Details of Housing IV

Look for relationship between Δ housing prices and RRD

- Concerns about reverse causality and confounding factors (e.g., land use regs, prod. shocks to clusters)
- Turn to IV that exploits varied exposure to regional housing cycles (Guren et al. '22)

$$P_{cdt} = \delta_c \bar{P}_{(-c)dt} + \psi_0 \hat{\beta}_{ct} + \psi_1 m_{cdt} + \phi_c t + D_c + \epsilon_{cdt} \quad \text{(Step 1)}$$

$$\hat{\beta}_{ct} = \gamma P_{cdt} + D_c + T_t + e_{ct} \quad \text{(Step 2)}$$

$$\mathbb{E}[\hat{\delta}_c \bar{P}_{(-c)dt} \times e_{ct}] = 0$$

Time-varying cousin to Saiz elasticity instrument (Saiz 2010; Mian, Rao, & Sufi '13)

- Saiz IV is not time-varying and can be correlated with city characteristics (Davidoff '16)
- Exploits system variation in exposure to demand shocks (Palmer '15)

Housing Price Effect on RRD

	All Cities			Cities with >200k		
	OLS (1)	IV (2)	Sort. (3)	OLS (4)	IV (5)	Sort. (6)
A. Estimates						
P_{cdt}	0.0655*** (0.0162)	0.0494* (0.0246)		0.0620*** (0.0150)	0.0524* (0.0262)	
$\rho_{ct}(P, \tau)$			-0.0500* (0.0220)			-0.0754 (0.0541)
N	1705	1705	1673	450	450	450
B. First Stage						
$\hat{\delta}_c \bar{P}_{(-c)dt}$		0.6140*** (0.1315)			0.6056*** (0.1331)	
F-stat, CD/KP		1245/21.8			347/20.7	

- RRD and housing prices: 10% price increase $\rightarrow$ RRD up by 0.5 log points ($\approx$0.07 SDs)

Housing Price Effect on RRD

	All Cities			Cities with >200k		
	OLS (1)	IV (2)	Sort. (3)	OLS (4)	IV (5)	Sort. (6)
A. Estimates						
P_{cdt}	0.0655*** (0.0162)	0.0494* (0.0246)		0.0620*** (0.0150)	0.0524* (0.0262)	
$\rho_{ct}(P, \tau)$			-0.0500* (0.0220)			-0.0754 (0.0541)
N	1705	1705	1673	450	450	450
B. First Stage						
$\hat{\delta}_c \bar{P}_{(-c)dt}$		0.6140*** (0.1315)			0.6056*** (0.1331)	
F-stat, CD/KP		1245/21.8			347/20.7	

- RRD and housing prices: 10% price increase $\rightarrow$ RRD up by 0.5 log points ($\approx$0.07 SDs)
- $\rho_{ct}(P, \tau)$: within-CZ correlation between tract-level housing price & travel time
- Increased sorting on job access $\rho_{ct}(P, \tau) < 0$ increased RRD

Counterfactual: Housing Prices

Given a plausible effect of housing prices on RDD, how big is it?

$$RRD = \Delta^{\text{Explained}}(z_c) + \Delta^{\text{Unexplained}}(z_c)$$
$$\Delta^{\text{Explained}}(z_c) = \sum p_c \gamma z_c = \gamma \bar{z}$$

Let $RRD' = \Delta^{\text{Explained}}(z_c') + \Delta^{\text{Unexplained}}(z_c)$, so $RRD' - RRD = \gamma(\bar{z}' - \bar{z})$.

Counterfactual: Rewind housing prices to 1980, but keep everything else as in 2019:

- Real average CZ log housing price increased 0.431 from 1980 to 2019 (11.99 to 12.42)

$$0.0494 \times 0.431 = 0.021$$

Housing prices return to 1980 levels would decrease RRD by 2.1 log points...

- Roughly 40–55% of its 2012–19 value

Summary

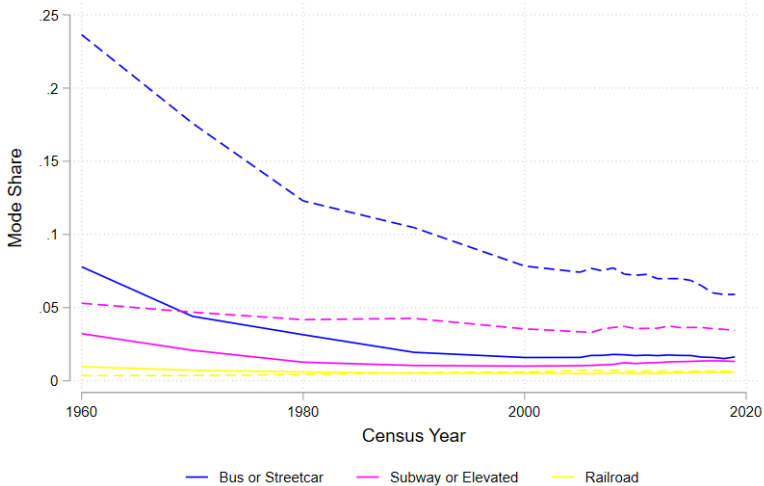
- Substantial—but incomplete—convergence in commute times by race since 1980
- Racialized difference, once systematic across the US, is now most present (i) in bigger cities for all commuters and (ii) for transit users and walkers everywhere
  - Accounting for job/income now increases difference.
  - Differences present across the income spectrum, but larger for lower-income workers
- Large cities contain ingredients of stratification associated with racialized difference
- Increasing housing prices in big, expensive, and congested cities exacerbate racialized difference today
  - Suburbanization trends of Black employment and residential location do not necessarily overlap spatially (Bartik & Mast '21; Kneebone & Holmes '15; Miller '18)

Questions that we have...

- How do existing patterns of residential segregation and place of work interact?
- Much wealth inequality is intergenerational, and this impacts residential location choice. How does this impact labor market access and outcomes?
- What policy interventions might be useful?
  - Lower housing price growth?
  - Better transit provision?

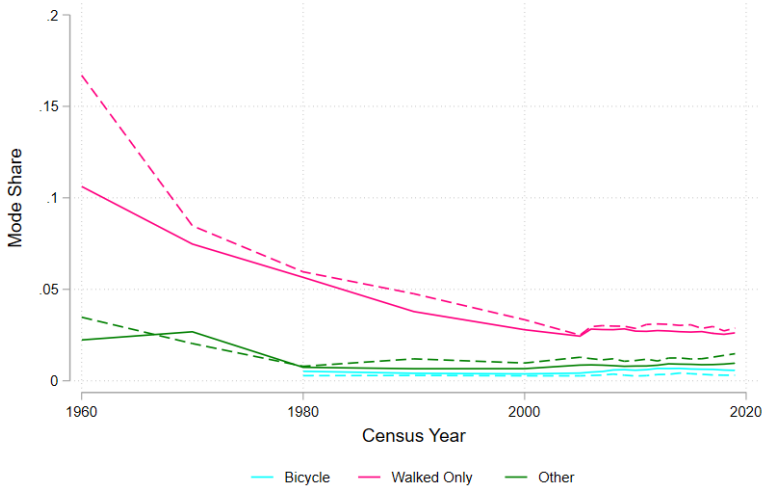
Thank you!

Mode Share - Transit



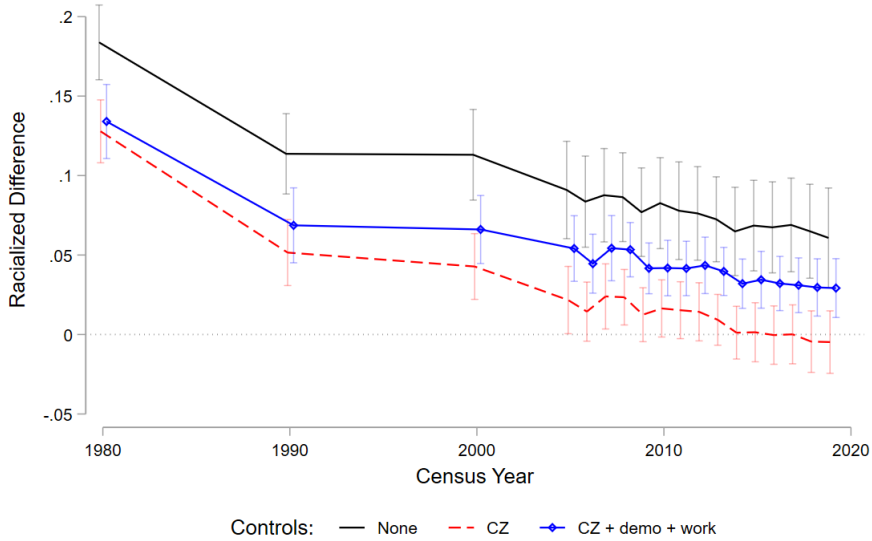
Return

Mode Share - Walk/Other

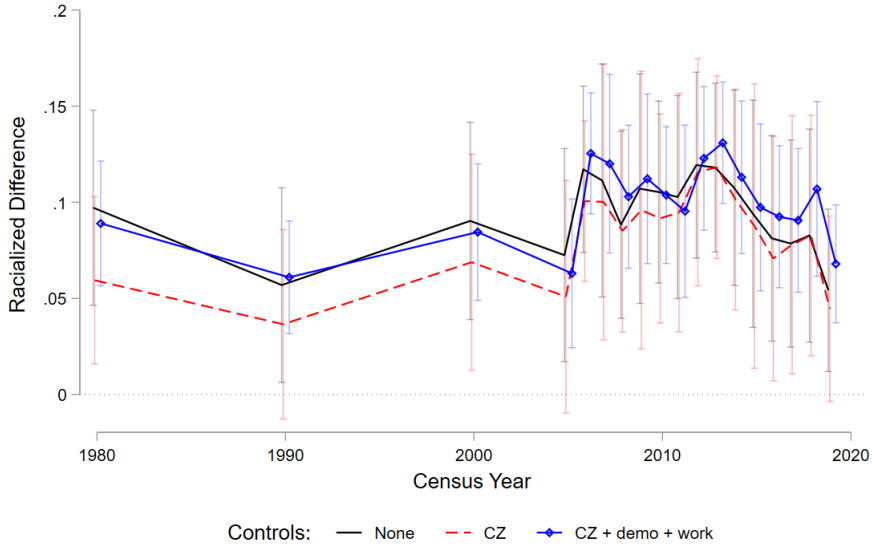


Return

Baseline Results - Car



Baseline Results - Bus



Baseline Results - Subway