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

PRELIMINARY; PLEASE DO NOT CITE WITHOUT PERMISSION.

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...”

- Quote from Dr. Martin Luther King Jr. about discrimination faced by Black bus riders, made during the Montgomery Bus Boycott (1955).
- Plessy v. Ferguson (1896), which legitimized ‘separate but equal’, was about segregation on trains.

$\Rightarrow$ Racialized difference in transportation is pervasive throughout US history

This paper: Provide a comprehensive accounting of racialized differences in commuting from 1980 to 2019

- Update prior literature in economics and sociology and study trend
- Decompose differences: mode, location, observables
- Study aggregate (city-level) determinants
- Suggest an interpretation

Summary of Findings

1. Black commuters face longer commutes than White commuters
   - Difference declining since 1980, but still persists (even conditionally)
2. Amongst transit users, the difference has not decreased or has increased
   - Suggests partial convergence due largely to car adoption by Black commuters
   - But there is still a small but significant difference for car commuters
3. Within-city res. location (PUMA) does not account for much of the difference
4. Difference largest at lower incomes, but are present at high incomes too
5. While differences have mostly shrunk since 1980, they persist in:
   - Large, congested cities that have experienced rapid house price growth.
   - IV to show high housing costs drive at least part of the residual difference
   - Evidence supports increasing spatial stratification on commute times in $$$ cities

Data

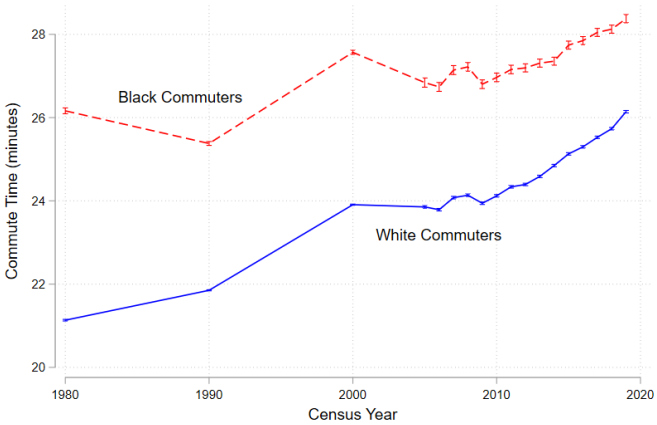
Census/ACS, 1980–2019, for all commuters

- Journey to Work questions ask about race and commute time/mode
- 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
- Extend back to 1960 for aggregate mode share

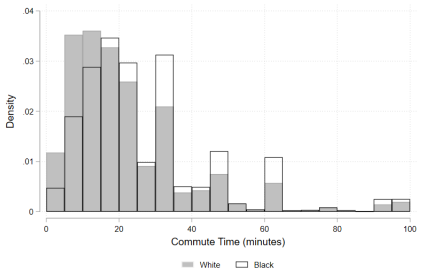
Additional sources

- NHGIS for finer geographic aggregates
- Zip Code Business Patterns for spatial dist. of work locations

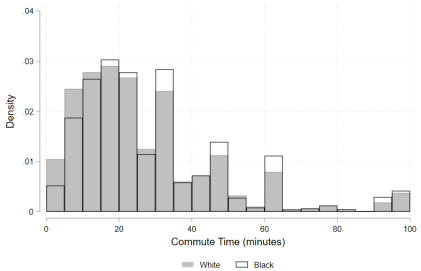
Aggregate Differences (Time)



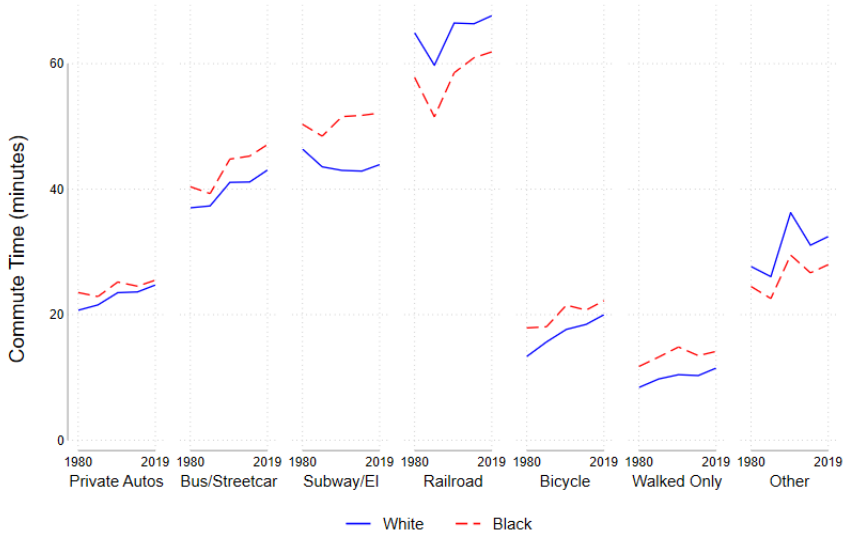
1980



2012–19



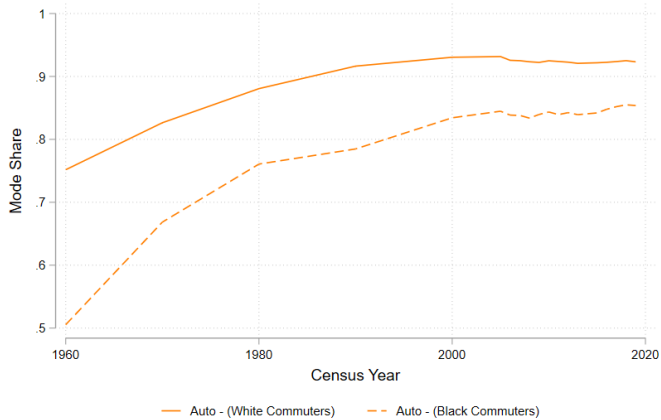
Aggregate Differences (Time by Mode)



Aggregate Differences (Mode)

Major increases in auto commuting

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



Analytic Framework

Use regression compatible decomposition approach (Fortin '08, Fortin et al. '11)

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

Group obs. covariates into categories:

- CZ
- Demographics/Educ
- Transportation Mode
- Job/Income

Permit time-varying coeffs

Can estimate β_{ct} (at city-level)!

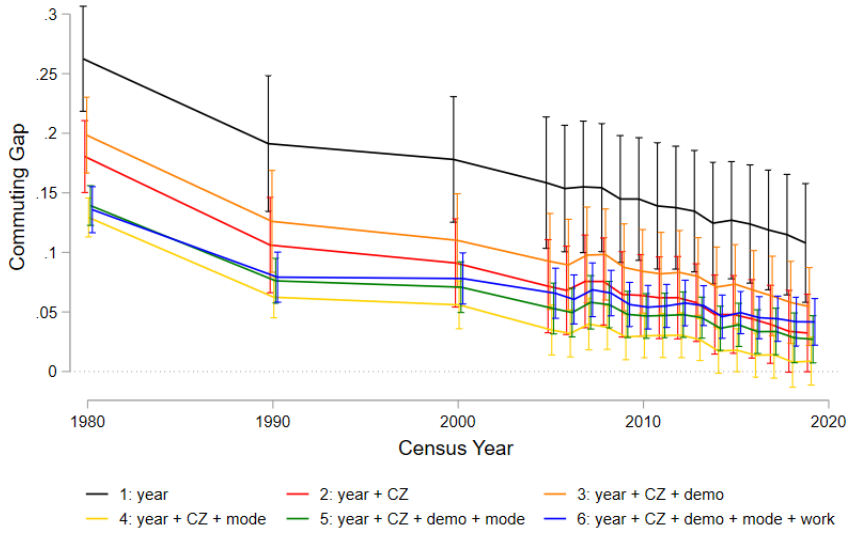
CAUTION: Intepretation

- What does it mean to control for e.g., income?
- Discrimination or structural racism could drive income differences
- Interpret as potential mechanisms

CAUTION: Selection

- Many margins of selection, esp: LFP, mode, and job

Baseline Results



Decomposition

Decompose along observable dimensions (non-sequentially a la Gelbach '16)

	$\Delta_{\{t\}}$	$\Delta^{\text{Unexplained}}_{\{t\}}$	$\Delta^{\text{Explained}}_{\{t\}}$			
			$\Delta^{\text{Demog.}}_{\{t\}}$	$\Delta^{\text{Tr. Mode}}_{\{t\}}$	$\Delta^{\text{Work/Inc.}}_{\{t\}}$	$\Delta^{\text{CZ}}_{\{t\}}$
$1[\text{Black}] \times t_{1980}$	0.263***	0.136***	-0.008***	0.073***	-0.001	0.062***
	(0.022)	(0.010)	(0.000)	(0.016)	(0.002)	(0.008)
		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***
	(0.029)	(0.011)	(0.000)	(0.018)	(0.002)	(0.009)
		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***
	(0.027)	(0.011)	(0.000)	(0.013)	(0.002)	(0.009)
		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***
	(0.027)	(0.010)	(0.000)	(0.014)	(0.002)	(0.009)
		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***
	(0.025)	(0.009)	(0.000)	(0.012)	(0.001)	(0.010)
		39.1%	-6.6%	32.5%	-15.4%	50.4%

$N = 46.8$ mil.

Decomposition

Contribution of **location (CZ)** to difference is ~ constant (6–7 log points)

- Makes up a larger share as unconditional difference shrinks

Demographics and **job/income** do not play a huge role

- Typically *increases* difference!
- **Job/income** playing a larger role over time

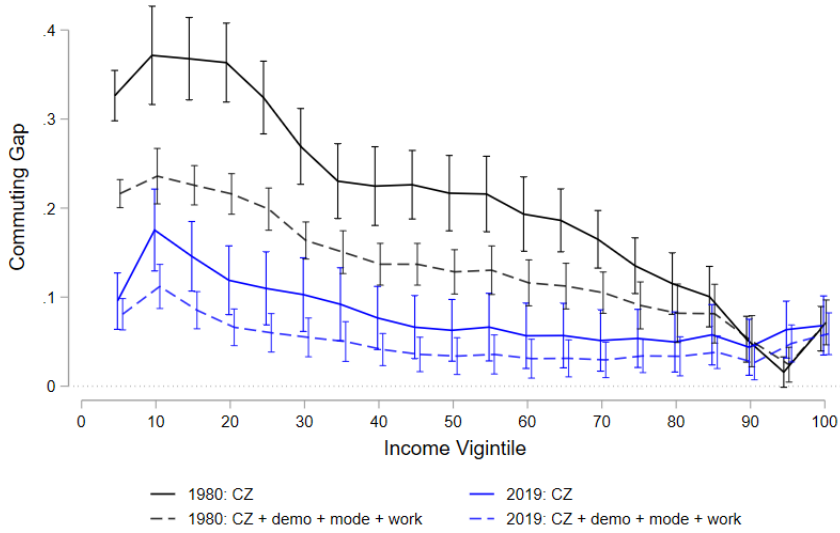
Transportation mode makes up a roughly constant share of difference

- One-quarter of decline is explained by partial convergence in mode share

Share of difference **unexplained** by our observables falling

- Still about 39% of unconditional difference

Heterogeneity by Income



Additional Aggregate Specs

Differences conditional on mode

- Car: Strongest evidence of partial convergence Car
- Bus & Subway: Difference flat or growing ($\sim$ 10 log points) Bus Subway

With PUMA FEs (2000 and later, might extend to 1990)

- Makes surprisingly little difference (except for subway)
- Meso-scale residential sorting doesn't explain difference

By city type

- Differences larger in bigger cities w/ heavy rail (even for auto)
- An exception: walkers, much larger differences in non-large non-transit CZs

City-level Heterogeneity

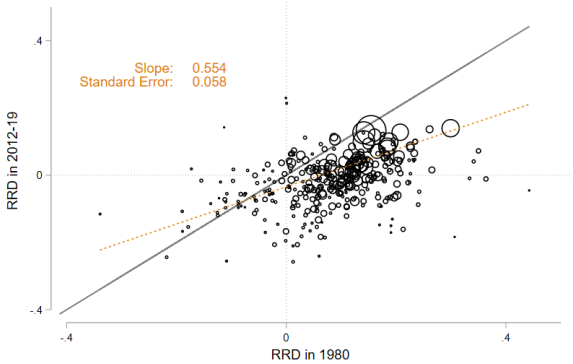
What drives city-level variation in this difference:

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

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

- RRD can be shown to contribute to $\Delta^{\text{Unexplained}}_{\{t\}}$
- Dealing with generated β and heteroskedasticity
  - Drop CZs with $<$1k commuters, CZs with $<$50 unique Black commuter Census respondents
  - Weight second stage by number unique Black commuter Census respondents (could also weight by robust variance estimate of $\hat{\beta}_{ct}$
  - Cluster SEs by CZ

Summary and Persistence



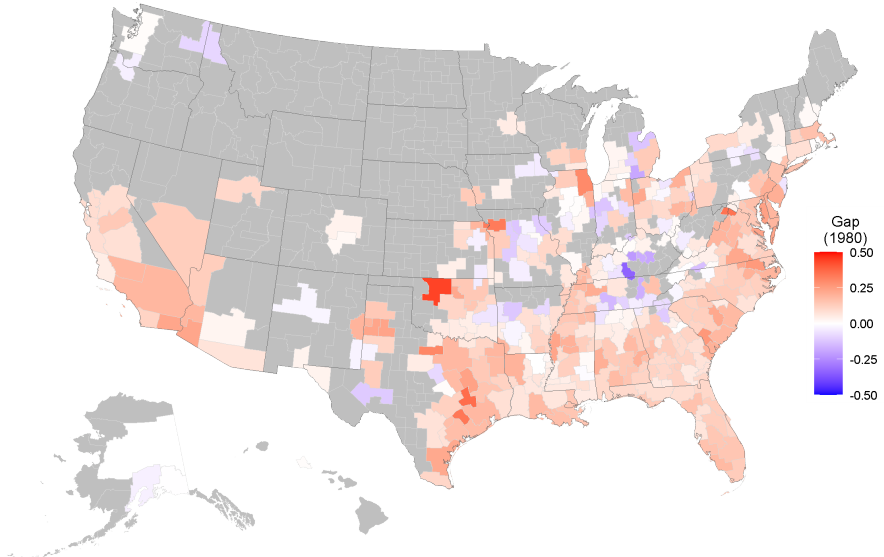
	Mean	SD	Min	Max
1980	0.131	0.072	-0.339	0.442
1990	0.070	0.072	-0.326	0.246
2000	0.068	0.077	-0.412	0.247
2005–11	0.053	0.073	-0.384	0.220
2012–19	0.036	0.070	-0.257	0.230

$N_t = 339$

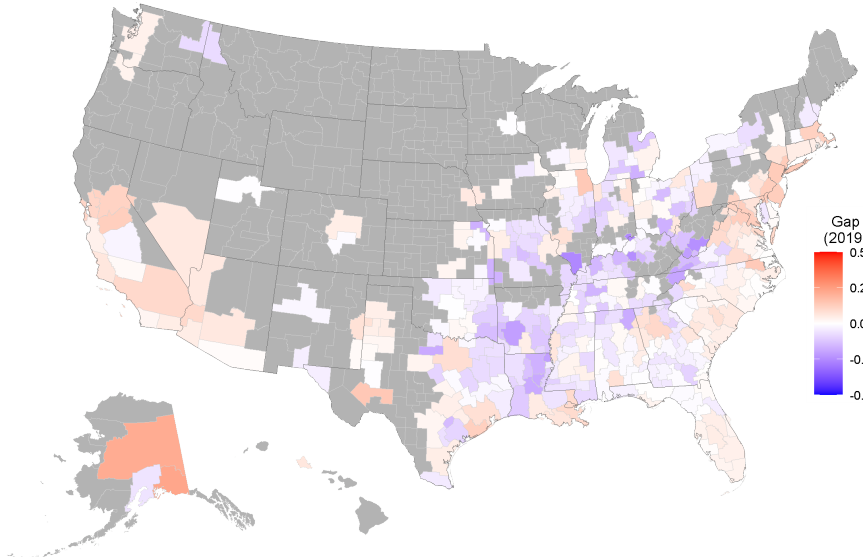
Declining mean, but not much decline in SD

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

Map (1980)



Map (2012–19)



Correlates

	1980		2000		2012–19	
	(1)	(2)	(3)	(4)	(5)	(6)
Ln(Population)	0.015***	0.022***	0.028***	0.034***	0.029***	0.032***
	(0.004)	(0.005)	(0.003)	(0.003)	(0.002)	(0.002)
% Black		0.318***		0.253***		0.123***
		(0.050)		(0.038)		(0.026)
Obs.	339	339	339	339	339	339
R-sq	0.131	0.264	0.380	0.475	0.545	0.581

- Population playing an increasingly large role in RRD
- Black share of population playing a smaller role

- Large (& often coastal cities) cities see much of the RRD $\Rightarrow$ Housing Costs?

Housing Prices

1. Access is a somewhat persistent 'second-nature' n'hood amenity (e.g., Cronon '91)
2. Big, expensive cities features lots of variation in job access
   - Geography lurking in the background (Saiz '10; Lee & Lin '18; Saiz & Wang '21)
3. Inelastic supply likely binds first in $$$ places $\rightarrow$ increased stratification (Van Nieuwerburgh & Weill '10; Guerrieri, Hartley, Hurst '13; Gyourko, Mayer, Sinai '13)

Housing Prices

1. Access is a somewhat persistent 'second-nature' n'hood amenity (e.g., Cronon '91)
2. Big, expensive cities features lots of variation in job access
   - Geography lurking in the background (Saiz '10; Lee & Lin '18; Saiz & Wang '21)
3. Inelastic supply likely binds first in $$$ places $\rightarrow$ increased stratification (Van Nieuwerburgh & Weill '10; Guerrieri, Hartley, Hurst '13; Gyourko, Mayer, Sinai '13)

Test link between rising housing prices and stratification

- Estimate panel (TWFE) model of RRD on housing prices
- Use Guren et al. '21-type IV (like Saiz elasticity) Description

$\Rightarrow$ High housing costs play a role at preserving RRD

	OLS		IV	
Ln(Hous. Pr.)	0.057**	0.058**	0.043+	0.044+
	(0.013)	(0.013)	(0.025)	(0.024)
Perc. Black		0.059		0.046
		(0.125)		(0.123)
First Stage			0.583**	0.573**
			(0.096)	(0.099)
F-stat (CD)			1070	1039
F-stat (KP)			37	34
Obs.	1695	1695	1695	1695

Other Aggregate Measures

Estimate additional correlations

- Cross section: with and without population control
- Panel: with TWFEs (CZ & year-bin), with and without population control

Suggestive evidence that

- Larger difference persists in transit-heavy, longer-commute (slower) CZs
- Urban form, segregation, daytime vs. nighttime population variables only matter insomuch as they are correlates of population
- Places with increasingly correlated (Expensive Housing, Short Commute) see higher RRD (supports stratification)

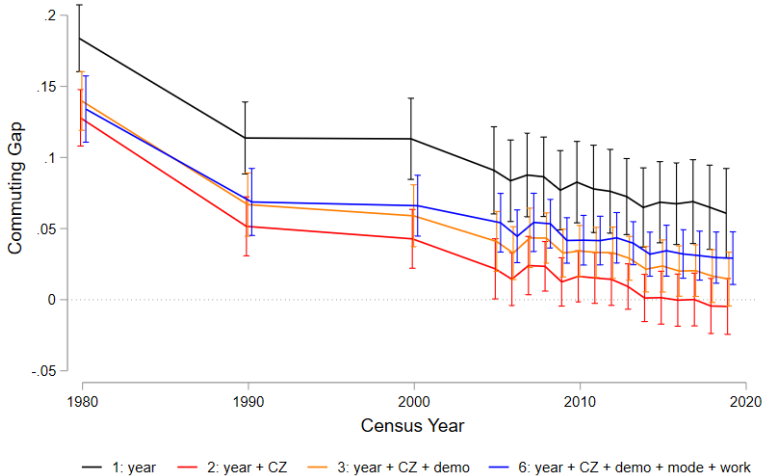
Table

To Do - Thoughts and Comments?!

1. Additional aggregate city correlates of RRD
   - Race-specific urban form measures
2. Stratify analysis by city size/history
   - E.g., Large vs. Medium CZs, Old vs. New CZs.
3. How important is selection?
   - Selection into employment (not commute times if not)
   - Selection in mode, etc.

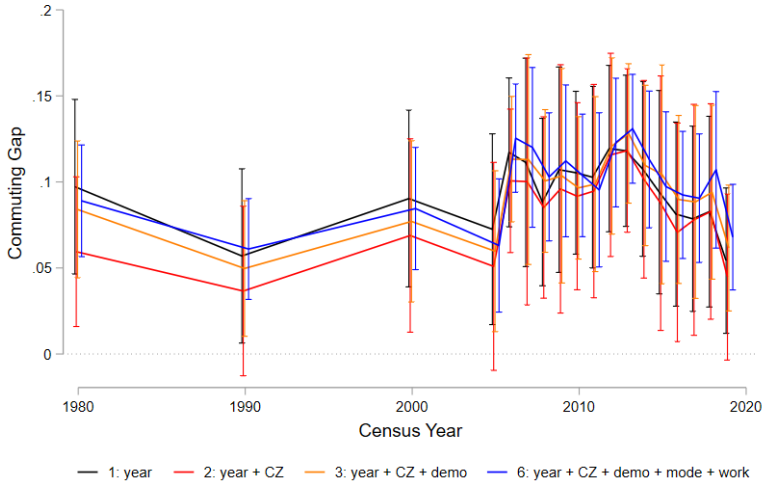
Thank you!

Baseline Results - Car



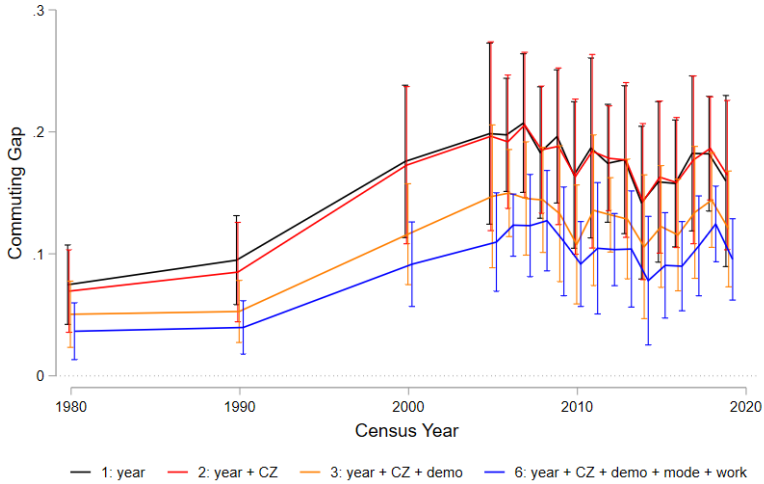
Return

Baseline Results - Bus



Return

Baseline Results - Subway



Return

Details of Housing IV

Confounders, e.g.: localized land use regs, prod. shocks to clustered industries

Estimate division-level variant of IV in Guren et al. '21

$$P_{cdt} = \delta_c \bar{P}_{(-c)dt} + \psi m_{cdt} + \phi_c t + D_c + \epsilon_{cdt}$$

- 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

Identification excludes the presence of any unobserved factor that

i. is correlated with regional house price movements, and
ii. differentially impacts the RRD according to housing price sensitivity

Return

Other Aggregate Measures – Results

Dependent Variable: RRD

	Aggregates		Commuting Chars.			Urban Form/Sorting				Comm. Strat.	
	Ln(Pop.)	% Black	% Transit	Ave. Car Time	Ave. Transit Time	Centr' ity	Dis- simi- larity	Hut- chens	Zip Empl. Gini	Corr(Time, Hous.)	%Hous. -%Time
Unconditional											
Measure	0.070***	0.005	0.289+	0.006**	0.001***	0.065	-0.069	-0.043	-0.001	-0.025*	-0.094*
	(0.018)	(0.128)	(0.150)	(0.002)	(0.000)	(0.142)	(0.048)	(0.067)	(0.045)	(0.010)	(0.047)
Controlling for Log Population											
Measure		0.037	0.287+	0.004*	0.001**	0.181	-0.021	0.012	0.009	-0.020*	-0.079+
		(0.099)	(0.156)	(0.002)	(0.000)	(0.121)	(0.040)	(0.057)	(0.040)	(0.009)	(0.046)

Return