Using Mobile Device Activity Data to Study Local Variation in Onsite Work

Katharine G. Abraham, Mohammad Ashoori, Aref Darzi, Nathalie Gonzalez-Prieto, John C. Haltiwanger, Aliakbar Kabiri, and Erkut Y. Ozbay

CRIW Conference on the Changing Nature of Work

March 6, 2025

Overview

- Use data on mobile device locations to track OnSite Work (OSW)
  - Millions of high frequency geocoded observations allow algorithmic identification of devices' home and work locations
  - Focus on pandemic and aftermath as a case study that also lets us evaluate strengths and limitations of device location data for tracking work activity
- Two approaches to studying variation across tracts (neighborhoods) in how OSW changed following the pandemic :
  - Conditional (longitudinal) analysis: Follow individuals identified as working onsite in February 2020 to construct tract-level OSW persistence estimates
    - Works well for short horizons; have produced estimates for May and August 2020
  - Unconditional analysis: Study ratio of tract-level OSW in September 2020, 2021 and 2022 to tract-level OSW in February 2020.
- Results show large and systematic variation across tracts in the evolution of OSW following pandemic's onset
  - Much of this variation within states, within cities and even within counties

Strengths and challenges of using mobile device location data to track changes in work locations

Strengths:

- High frequency location information for millions of devices
  - Through algorithmic identification of devices' home and work locations, able to observe changes in OSW prevalence for detailed geographies

Challenges:

- Even with millions of observations, device location data only available for a subset of the population
  - Better suited for tracking *changes* in OSW than *level* of OSW
- Landscape for access to device location data is changing; data may be less available in the future

Conditional analysis: Track OSW from February 2020 to May and August 2020

- Basic idea: Study evolution of OSW in a sample for which February 2020 home and work locations can be confidently identified
- Home and work locations identified algorithmically
  - Home location: Most frequently observed device location, so long as observed for at least 60 hours and on at least 14 distinct days.
  - Work location: Second most frequently observed device location, so long as observed for at least 60 hours and during two distinct weeks.
    - Should capture full time OSW and substantial part time OSW (e.g., work on a hybrid schedule with several days a week onsite)
- Starting with 4.2 million devices with February 2020 home and work locations, retain those with identified May (August) 2020 home locations, then attempt to identify their work locations
  - May (August) 2020 samples weighted to account for attrition

Conditional analysis: Track OSW from February 2020 to May and August 2020 (continued)

- Calculate weighted share of devices in each tract with identified OSW in May (August) 2020
- Follow same procedures with data for February, May and August 2019
  - Think of 2019 estimates as capturing changes in OSW due to normal labor market turnover
- Focus on difference between May(Aug) 2020 and May(Aug) 2019 tract level estimates
  - Think of difference between the 2020 and 2019 estimates as capturing pandemic effects

Unconditional analysis: Track ratio of OSW in Sept. 2020, 2021 and 2022 to OSW in Feb. 2020

- Algorithms for identifying home and work locations similar to those for conditional analysis
  - Home location: Most frequently observed device location, so long as observed for at least 3 days, on at least half of all observed days, and for average of 2+ hours on those days.
  - Work location: Location other than home location observed most frequently during daytime weekday hours, so long as satisfies same frequency conditions applied to weekdays
  - 35+ million usable devices (devices with an identified home location) in every month
  - Devices average substantial hours at both home and work locations
    - In February 2020, mean of 187 home hours and 51 work hours
- OSW share in month: Number of devices with work location divided by number of devices with home location
- Outcome of interest: *Ratio* of OSW share in September 2020, 2021 or 2022 to OSW share in February 2020
  - One minus ratio approximates the proportional decline in OSW among the employed so long as 1) patterns of device use reasonably stable and 2) changes in EPOP ratios small relative to changes in likelihood of OSW

External validity checks lend credibility to measures based on mobile device location data

Conditional Analysis

	Among those working onsite in February 2020, percent OSW as of:	
	May 2020	August 2020
MTI/CATT Lab	53.0	62.6
RPS	52.9	60.5
Difference	0.1	2.1

Unconditional Analysis

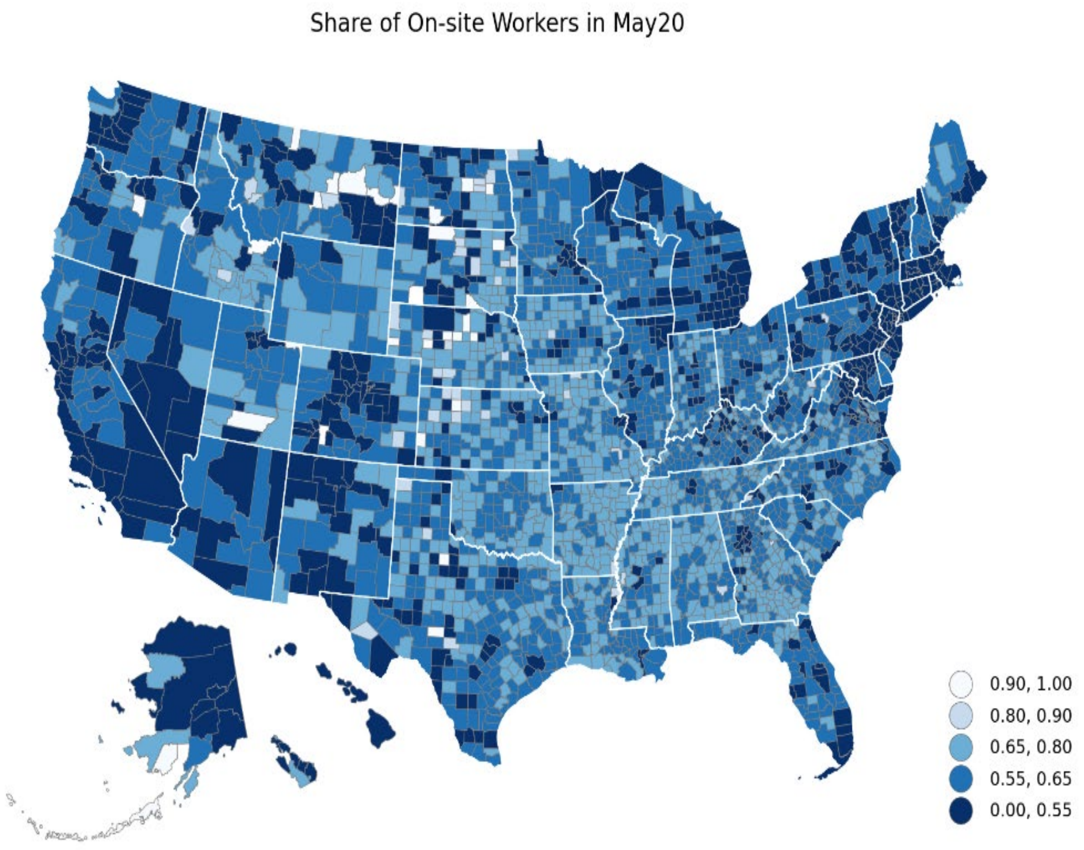
	Ratio of OSW in later month to OSW in February 2020		
	September 2020	September 2021	September 2022
MTI/CATT Lab	0.75	0.84	0.91
RPS	0.77	na	na
Difference	-0.02	na	na
	Ratio of OSW in later year to OSW in 2019		
		2021	2022
ACS	na	0.83	0.89
Difference	na	-0.01	-0.02

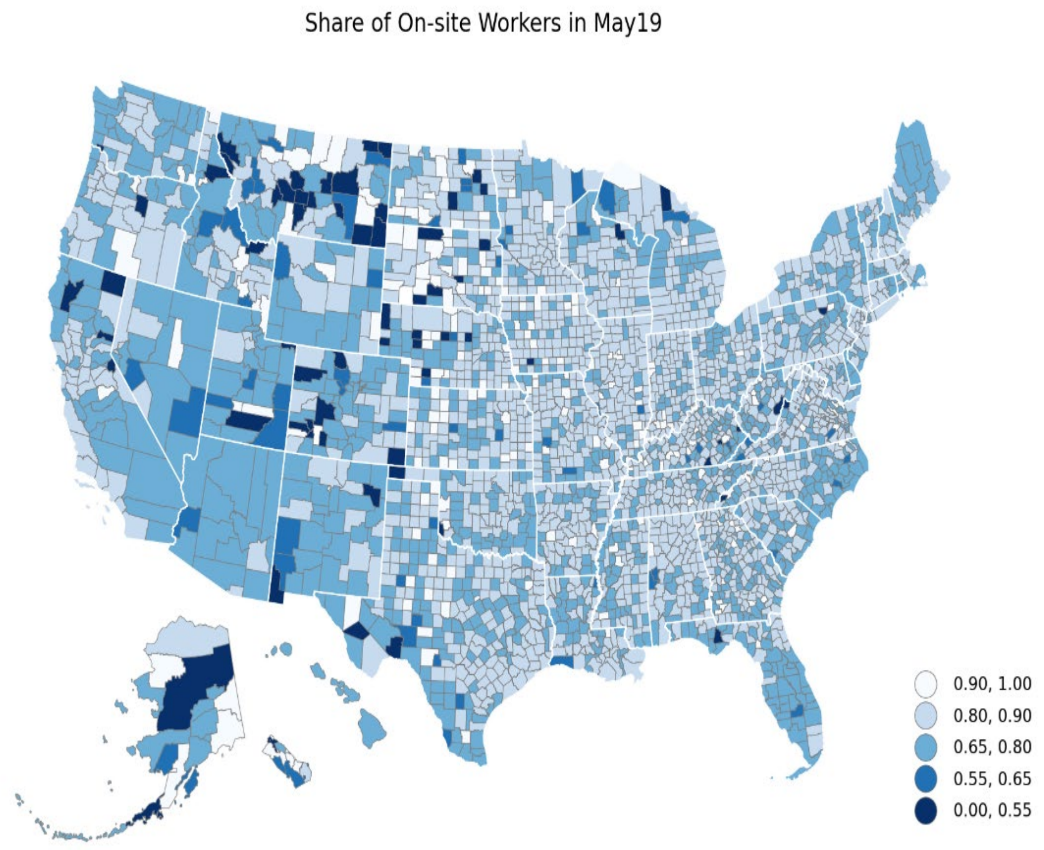
MTI/CATT Lab=Mobile device data; RPS=RealTime Population Survey; ACS=American Community Survey

Note: MTI/CATT Lab and survey estimates also highly correlated at Census Division Level.

Fraction of February 2020 OSW with May 2020 OSW

Fraction of February 2019 OSW with May 2019 OSW

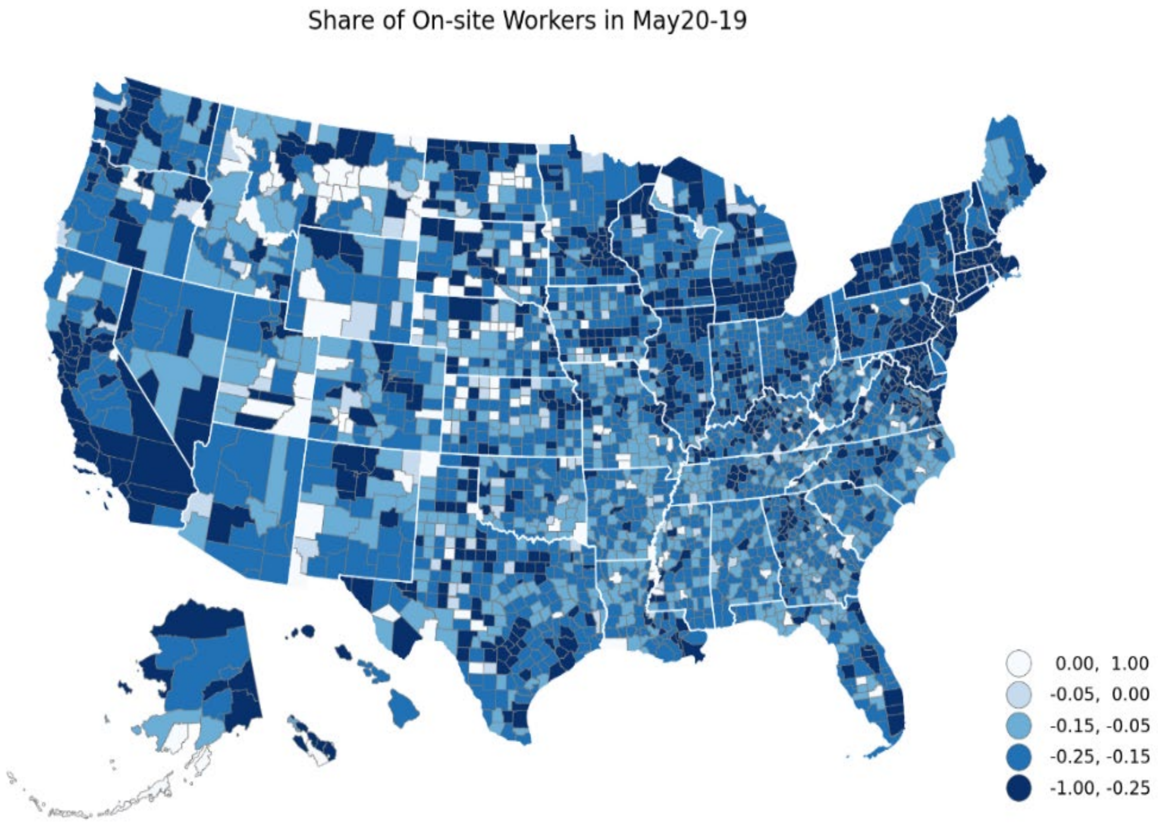


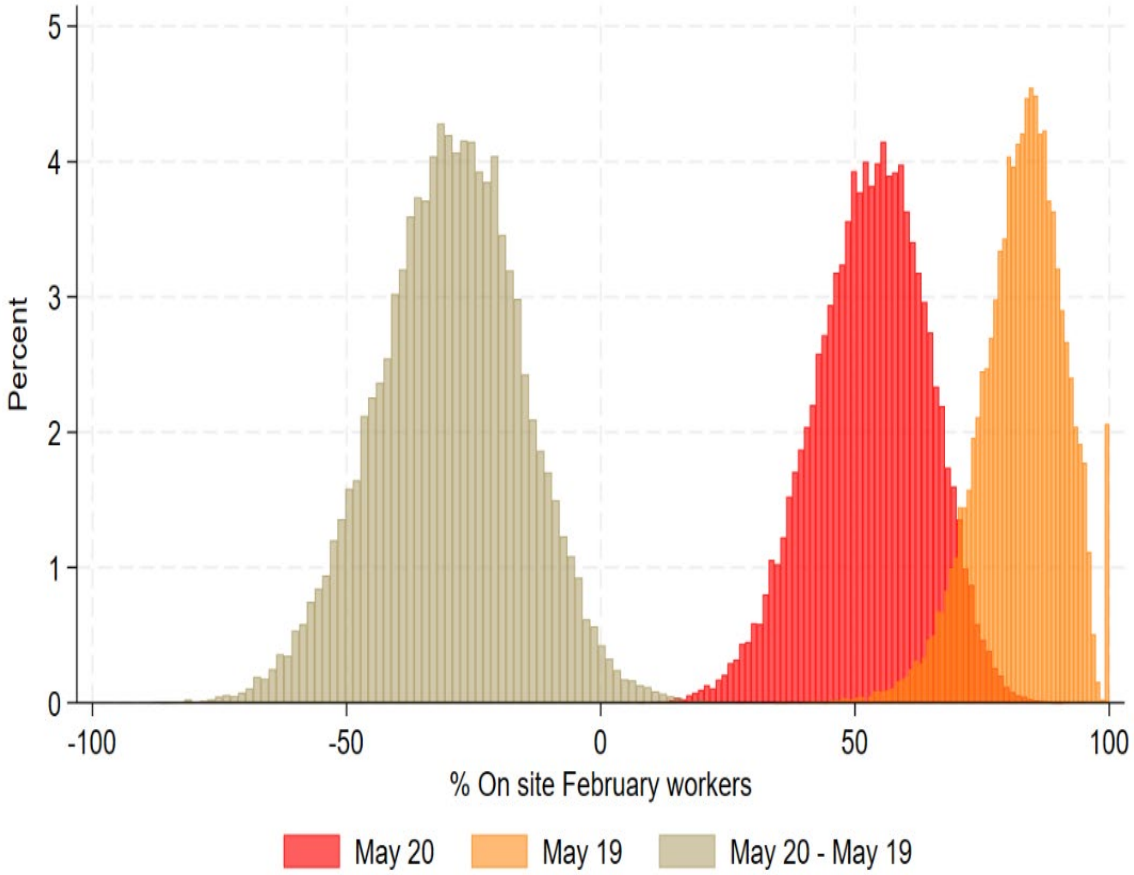


Note: For readability, estimates in maps displayed at county level, not tract level.

Difference in share of workers persisting in onsite work from February, May 2020 minus May 2019

Tract-level distribution of workers persisting in OSW, May 2020, May 2019 and May 20-May 19 difference



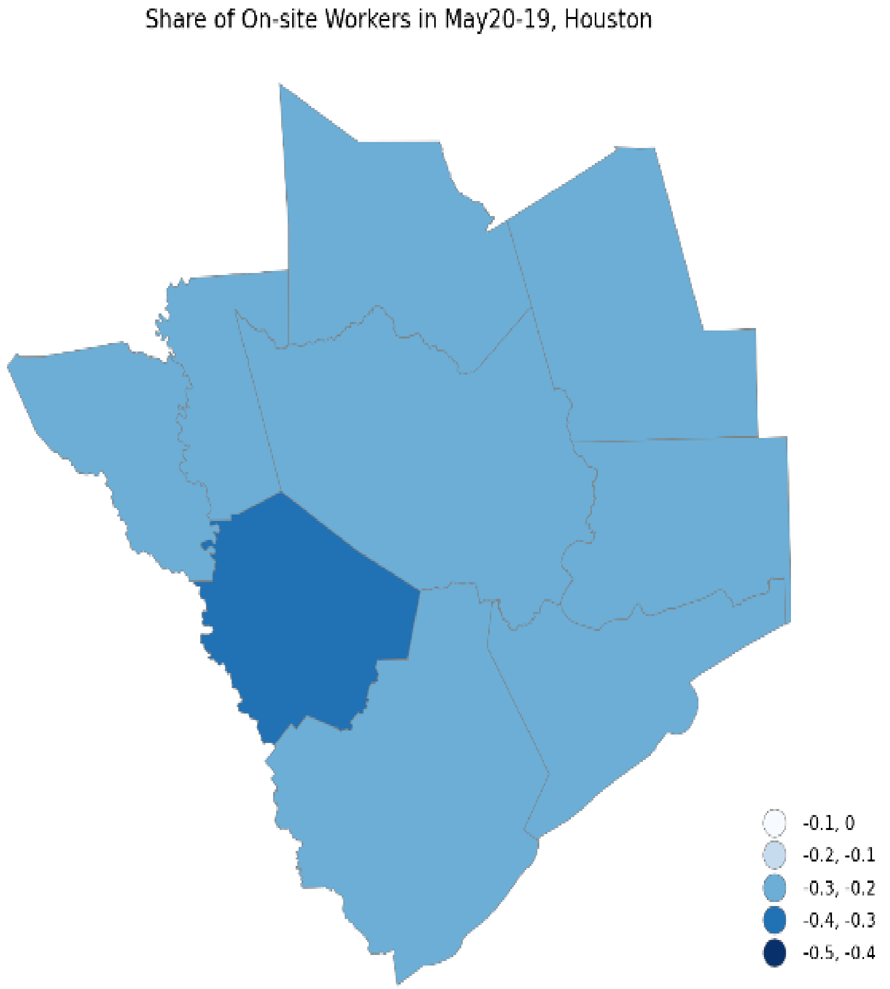


Note: For readability, estimates in map displayed at county level, not tract level.

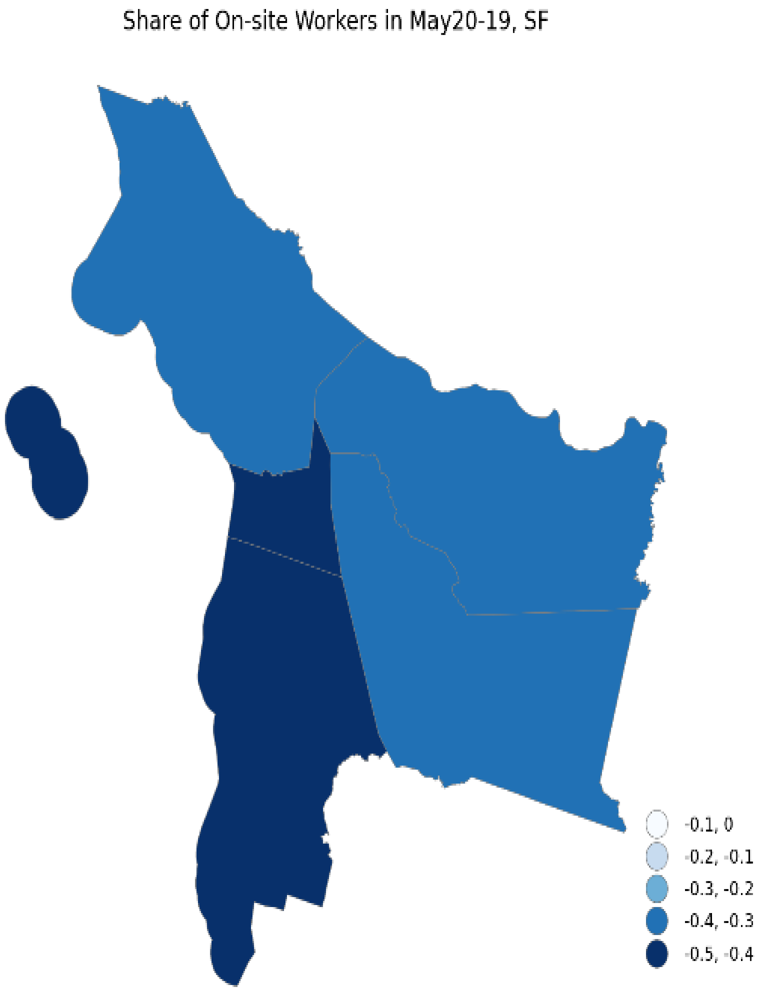
Mean (employment) weighted May 20-May 19 difference -30% (s.d. 14 pp). Within-county variance 64% of overall cross-tract variance.

OSW change from February to May 2020 minus OSW change from February to May 2019

Houston CBSA



San Francisco CBSA



Note: Estimates in maps displayed at county level

Percent of Tract-level Variance in 2020 minus 2019 Onsite Work (OSW) Persistence in Later Months Among Individuals with OSW in February Explained by Various Factors

Explanatory variables	May 2020 minus May 2019	August 2020 minus August 2019
Share of population:		
Age 25-64	-0.3	-0.1
Age 65 plus	0.0	-0.1
White, non-Hispanic	0.0	0.3
College graduate	0.6	0.2
ln(mean household income)	14.4	4.7
Share commute public trans.	0.0	0.1
Share commute 30+ mins.	1.6	1.7
Rural (yes/no)	0.6	0.2
Share Trump vote in 2016	1.2	0.1
May 2020 state lockdown	3.6	0.4
May 2020 local lockdown	0.3	0.1
May 2020 cum COVID deaths	2.5	-0.3
Industry mix	12.5	9.9
Occupation mix	7.8	6.9
Residual	55.4	75.7
Dep. var. mean	-29.6	-15.8
Dep. var. standard deviation	(14.4)	(14.7)

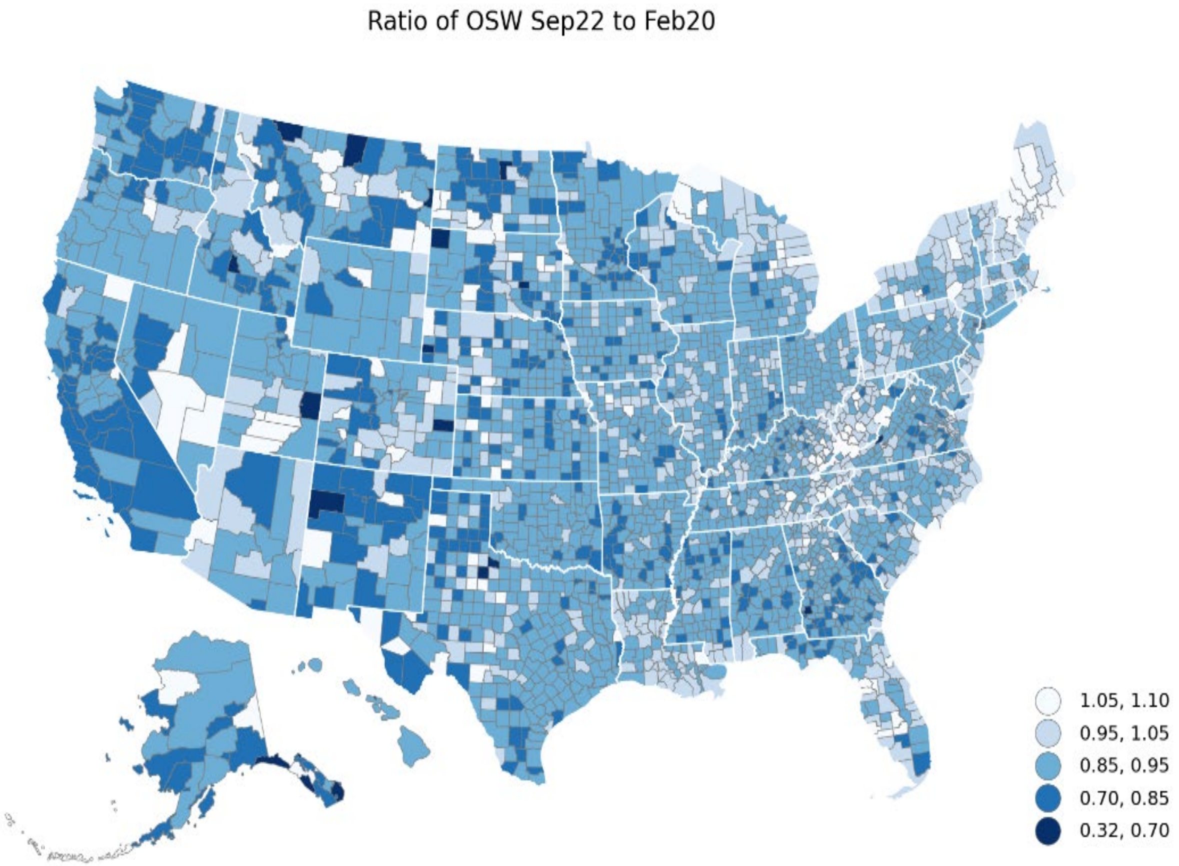
Covariates from 2015-2019 ACS; USDA; 2016 Trump vote share; COVID variables

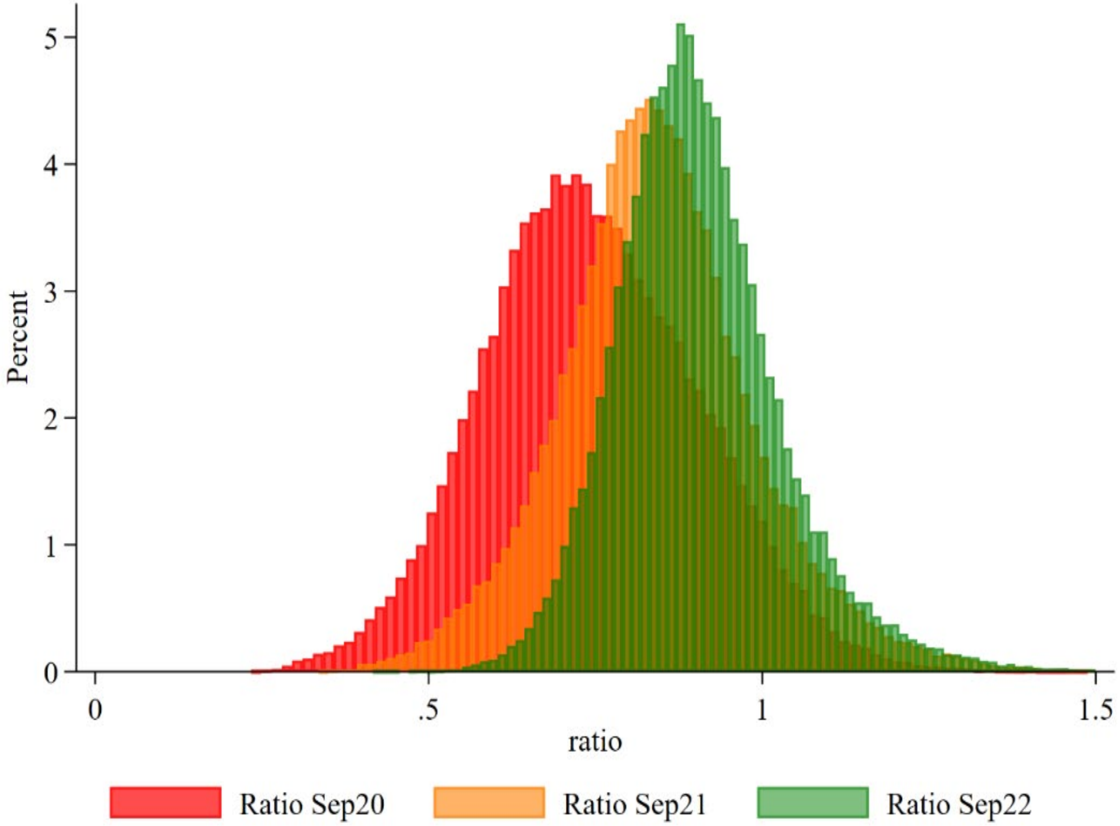
Model accounts for ~45% of variation in May20-May19 difference, ~25% of variation in Aug 20-Aug19 difference

Most important covariates: Household income, industry mix, occupation mix

Variation mainly within state and county, not across states and counties

Unconditional analysis: OSW recovering by September 2022 but still varied substantially across counties and tracts

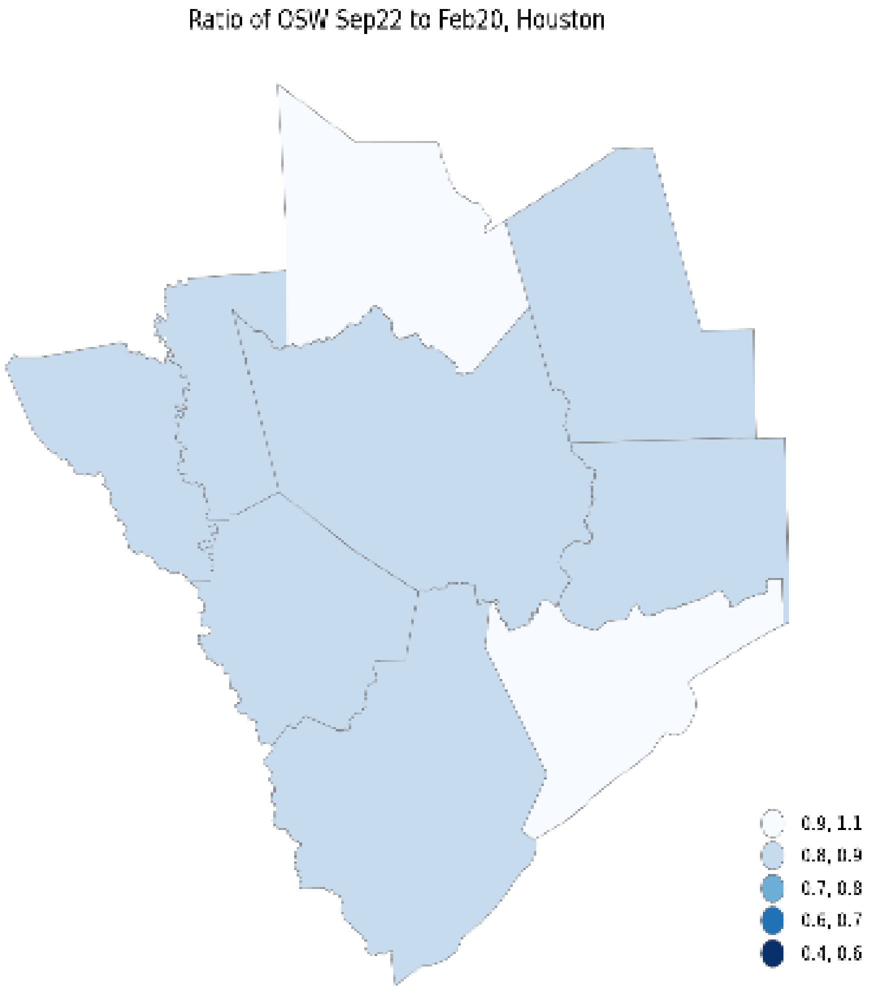


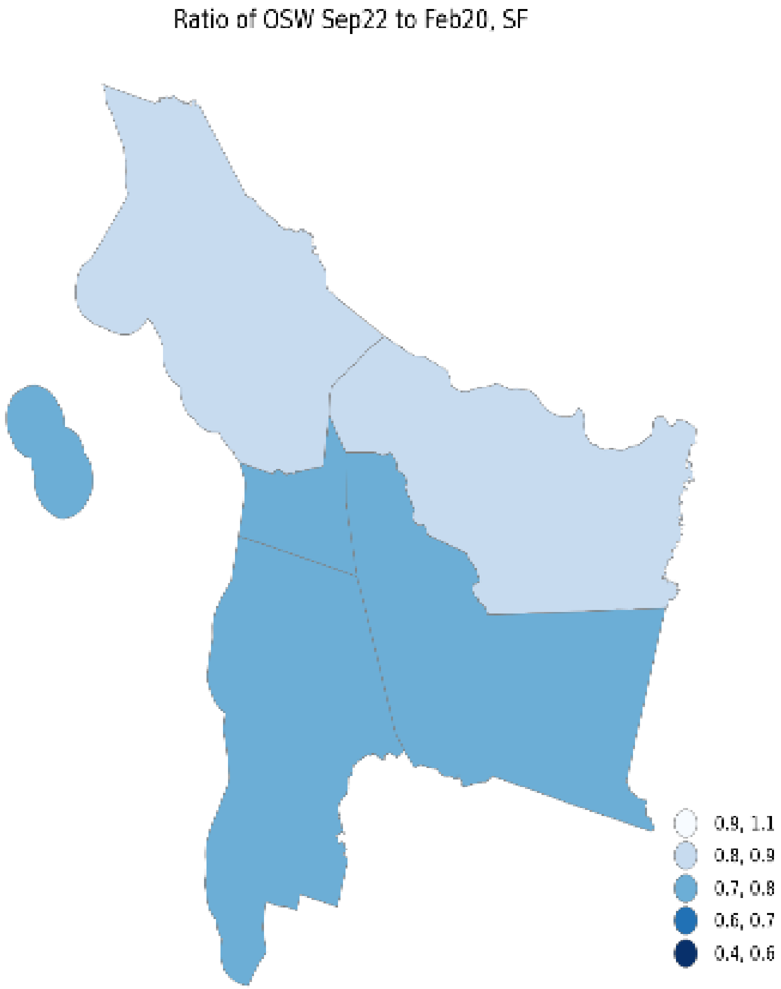


Note: For readability, estimates in map displayed at county level, not tract level.

Even in Sept. 2022, cross-tract variation remained large, with ~80% of (employment weighted) variation within counties rather than across counties.

Houston OSW ratio substantially above San Francisco OSW ratio in September 2022, but mostly well below 1.0 in counties in both cities





Note: Estimates in maps displayed at county level

Percent of Variance in Ratio of Prevalence of Onsite Work (OSW) in Later Months to Prevalence of OSW in February 2020 Explained by Various Factors

	Ratio of OSW in month to OSW in February 2020		
Explanatory variables	September 2020	September 2021	September 2022
Share of population:			
Age 25-64	-1.4	0.7	0.8
Age 65 plus	1.4	1.1	2.9
White, non-Hispanic	3.0	4.8	4.6
College graduate	5.2	-0.1	0.0
ln(mean household income)	9.5	5.2	0.3
Share commute public trans.	0.3	0.1	1.2
Share commute 30+ mins.	1.3	0.7	-0.3
Rural (yes/no)	4.4	0.5	-0.2
Share Trump vote in 2016	12.3	-0.9	1.6
May 2020 state lockdown	5.9	2.6	0.5
May 2020 local lockdown	0.1	0.0	0.0
May 2020 cum COVID deaths	-0.9	0.0	0.1
Industry mix	11.4	5.3	2.8
Occupation mix	7.3	1.3	1.0
Residual	40.4	78.6	84.9
Dep. var. mean	0.75	0.84	0.91
Dep. var. standard deviation	(0.16)	(0.15)	(0.13)

As in Aug 20-Aug 19 conditional analysis, income, industry mix and occupation mix explain sizable shares of variance in September 2020 ratio

2016 Trump vote share also important

Covariates account for ~60% of variation in September 2020 ratio but only ~15% of variation in September 2022 ratio

Neighborhood-level variation important for analysis and policy

- Some neighborhoods have much higher fractions of people continuing OSW with associated demand for local transportation services
- Other neighborhoods have much higher fractions of people working from home with associated demand for other local services during the work week
- One size fits all approach within a CBSA or even a county will not work

Taking stock

- Naturally occurring mobile device location data provide unique perspective on patterns of OSW (and other aspects of mobility):
  - Thanks to high frequency observations for millions of devices, have documented enormous cross-tract variation in OSW not observable on a timely basis using data from other sources
  - Most household surveys produce data only for large geographies
  - ACS provides tract level data, but only for rolling 5-year windows
  - Job postings data (e.g. Hansen et al. 2023) provide insights about changes in remote work prevalence, but refer to job locations, not home locations, and only for new hires
- Changing landscape may make it more difficult for future analysts to access large mobile device samples of the sort we have analyzed
  - Barring access issues, mobility data could in principle provide near-real-time estimates of how a specific event (pandemic, natural disaster, change in policy environment) is affecting the prevalence of onsite work