Ride-Sharing the Wealth – Effects of Uber and Lyft on Jobs, Wages and Economic Growth

SUPPLEMENTARY INFORMATION (SI)

Adam Koling
Saïd Business School
University of Oxford, Oxford, United Kingdom, OX1 1HP
Email: adam.koling@sbs.ox.ac.uk
Corresponding Author

Daniel Armanios
Saïd Business School
University of Oxford, Oxford, United Kingdom, OX1 1HP
Email: daniel.armanios@sbs.ox.ac.uk

Jeremy Michalek
Department of Engineering and Public Policy
Department of Mechanical Engineering
Department of Civil and Environmental Engineering
Carnegie Mellon University, Pittsburgh, Pennsylvania, 15213
Email: jmichalek@cmu.edu

Connor Forsythe
Department of Mechanical Engineering
Carnegie Mellon University, Pittsburgh, Pennsylvania, 15213
Email: connorforsythe@cmu.edu

Akshaya Jha
Heinz College of Information Systems and Public Policy
Carnegie Mellon University, Pittsburgh, Pennsylvania, 15213
Email: akshayaj@andrew.cmu.edu

Contents

1 TNC Regions **3**

2 Dataset **4**

3 Public Statements by Uber and Lyft Leadership **4**
3.1 Uber and Lyft claims about overall employment, wages, and GDP 4
3.2 Uber and Lyft claims about employment and wages in unstable jobs or industries . 5

4 Qualitative Rationale for Alternative (Non-TWFE) Estimators **6**

5 Goodman-Bacon Diagnostic **8**

6 Aggregating Group-Time Average Treatment Effects (Cohort, Dynamic) **8**

7 Additional Estimators (Sun and Abraham, Stacked Regression) **10**

8 Event Studies for Additional Estimators **10**

9 Data Availability and Extended Event Studies **16**

10 Identification Assumptions **16**
10.1 Callaway & Sant'Anna Assumptions 17
10.2 Correlated entry of other gig economy platforms 17

TNC REGIONS

Table 1 lists each TNC region together with its associated metropolitan or micropolitan statistical area name.

TABLE 1: Corresponding TNC Regions and Metropolitan/Micropolitan Statistical Areas

TNC Region	MSA Name	TNC Region (Continued)	MSA Name (Continued)
Abilene	Abilene, TX	Los Angeles	Los Angeles-Long Beach-Santa Ana, CA
Akron	Akron, OH	Louisville	Louisville/Jefferson County, KY-IN
Albuquerque	Albuquerque, NM	Lubbock	Lubbock, TX
Lehigh Valley	Allentown-Bethlehem-Easton, PA-NJ	Macon	Macon, GA
Amarillo	Amarillo, TX	Mankato	Mankato-North Mankato, MN
Ann Arbor	Ann Arbor, MI	Rio Grande Valley	McAllen-Edinburg-Mission, TX
Asheville, NC	Asheville, NC	Southern Oregon	Medford, OR
Athens	Athens-Clarke County, GA	Memphis	Memphis, TN-MS-AR
Atlanta	Atlanta-Sandy Springs-Marietta, GA	Miami	Miami-Fort Lauderdale-Pompano Beach, FL
Augusta	Augusta-Richmond County, GA-SC	NW Indiana	Michigan City-La Porte, IN
Austin	Austin-Round Rock-San Marcos, TX	Midland-Odessa	Midland, TX
Bakersfield	Bakersfield-Delano, CA	Milwaukee	Milwaukee-Waukesha-West Allis, WI
Baltimore-Maryland	Baltimore-Towson, MD	Minneapolis - St. Paul	Minneapolis-St. Paul-Bloomington, MN-WI
Baton Rouge	Baton Rouge, LA	Missoula	Missoula, MT
Bellingham	Bellingham, WA	Modesto	Modesto, CA
Central Oregon	Bend, OR	Eastern WV	Morgantown, WV
Billings	Billings, MT	Nacogdoches	Nacogdoches, TX
Bismarck	Bismarck, ND	Nashville	Nashville-Davidson–Murfreesboro–Franklin, TN
Boise	Boise City-Nampa, ID	New Orleans	New Orleans-Metairie-Kenner, LA
Boone	Boone, NC	New York City	York-Northern New Jersey-Long Island, NY-NJ-PA
Bozeman	Bozeman, MT	Sarasota	North Port-Bradenton-Sarasota, FL
Coastal Georgia	Brunswick, GA	Ocala, FL	Ocala, FL
Fort Myers-Naples	Cape Coral-Fort Myers, FL	Oklahoma City	Oklahoma City, OK
Cedar Rapids	Cedar Rapids, IA	Olympia	Olympia, WA
Champaign	Champaign-Urbana, IL	Omaha	Omaha-Council Bluffs, NE-IA
Charleston, SC	Charleston-North Charleston-Summerville, SC	Orlando	Orlando-Kissimmee-Sanford, FL
Chattanooga	Chattanooga, TN-GA	Ventura	Oxnard-Thousand Oaks-Ventura, CA
Chicago	Chicago-Joliet-Naperville, IL-IN-WI	Pensacola, FL	Pensacola-Ferry Pass-Brent, FL
Cincinnati	Cincinnati-Middletown, OH-KY-IN	Peoria, IL	Peoria, IL
Cleveland	Cleveland-Elyria-Mentor, OH	Philadelphia	Philadelphia-Camden-Wilmington, PA-NJ-DE-MD
Coeur D'Alene	Coeur d'Alene, ID	Pierre	Pierre, SD
College Station	College Station-Bryan, TX	Pittsburgh	Pittsburgh, PA
Colorado Springs	Colorado Springs, CO	Portland, ME	Portland-South Portland-Biddeford, ME
Columbia, SC	Columbia, SC	Portland	Portland-Vancouver-Hillsboro, OR-WA
Columbus, GA	Columbus, GA-AL	Raleigh-Durham	Raleigh-Cary, NC
Columbus	Columbus, OH	Rapid City	Rapid City, SD
Corpus Christi	Corpus Christi, TX	Reading, PA	Reading, PA
Dallas-Fort Worth	Dallas-Fort Worth-Arlington, TX	Reno	Reno-Sparks, NV
Quad Cities	Davenport-Moline-Rock Island, IA-IL	Richmond	Richmond, VA
Denver	Denver-Aurora-Broomfield, CO	Inland Empire	Riverside-San Bernardino-Ontario, CA
Des Moines	Des Moines-West Des Moines, IA	Roanoke-Blacksburg	Roanoke, VA
Detroit	Detroit-Warren-Livonia, MI	Rochester, MN	Rochester, MN
Dickinson	Dickinson, ND	Rockford	Rockford, IL
DuBois	DuBois, PA	Roswell	Roswell, NM
Dubuque	Dubuque, IA	Sacramento	Sacramento–Arden-Arcade–Roseville, CA
Duluth	Duluth, MN-WI	Tri-Cities, MI	Saginaw-Saginaw Township North, MI
Eagle Pass	Eagle Pass, TX	St Cloud	St. Cloud, MN
Eau Claire	Eau Claire, WI	Southern Utah	St. George, UT
El Paso	El Paso, TX	St Louis	St. Louis, MO-IL
Erie	Erie, PA	Willamette Valley	Salem, OR
Eugene, OR	Eugene-Springfield, OR	Salt Lake City	Salt Lake City, UT
Fargo - Moorhead	Fargo, ND-MN	San Angelo	San Angelo, TX
Flint	Flint, MI	San Antonio	San Antonio-New Braunfels, TX
Florence, SC	Florence, SC	San Diego	San Diego-Carlsbad-San Marcos, CA
Fort Collins	Fort Collins-Loveland, CO	San Francisco Bay Area	San Francisco-Oakland-Fremont, CA
Fort Wayne	Fort Wayne, IN	San Luis Obispo	San Luis Obispo-Paso Robles, CA
Fresno	Fresno, CA	Santa Fe	Santa Fe, NM
Gallup	Gallup, NM	Savannah-Hilton Head	Savannah, GA
Grand Forks	Grand Forks, ND-MN	Wilkes-Barre Scranton	Scranton–Wilkes-Barre, PA
Green Bay	Green Bay, WI	Seattle	Seattle-Tacoma-Bellevue, WA
Piedmont Triad	Greensboro-High Point, NC	Sioux Falls	Sioux Falls, SD
Harrisburg	Harrisburg-Carlisle, PA	South Bend	South Bend-Mishawaka, IN-MI
Big Island	Hilo, HI	Springfield, IL	Springfield, IL
Houston	Houston-Sugar Land-Baytown, TX	Springfield, Mo	Springfield, MO
Indianapolis	Indianapolis-Carmel, IN	State College	State College, PA
Iowa City	Iowa City, IA	Stillwater	Stillwater, OK
Jacksonville	Jacksonville, FL	Tallahassee	Tallahassee, FL
Johnstown-Altoona	Johnstown, PA	Tampa Bay	Tampa-St. Petersburg-Clearwater, FL
Maui	Kahului-Wailuku, HI	Taos	Taos, NM
Kansas City	Kansas City, MO-KS	Texarkana	Texarkana, TX-Texarkana, AR
Kauai	Kapaa, HI	Topeka	Topeka, KS
Florida Keys	Key West, FL	Traverse City	Traverse City, MI
Outer Banks, NC	Kill Devil Hills, NC	Tulsa	Tulsa, OK
Killeen	Killeen-Temple-Fort Hood, TX	Tyler	Tyler, TX
Tri-Cities	Kingsport-Bristol-Bristol, TN-VA	Honolulu	Urban Honolulu, HI
La Crosse	La Crosse, WI-MN	Hampton Roads	Virginia Beach-Norfolk-Newport News, VA-NC
Lancaster, PA	Lancaster, PA	Washington D.C.	Washington-Arlington-Alexandria, DC-VA-MD-WV
Laredo	Laredo, TX	Waterloo-Cedar Falls	Waterloo-Cedar Falls, IA
Las Cruces	Las Cruces, NM	Wichita	Wichita, KS
Las Vegas	Las Vegas-Paradise, NV	Wichita Falls	Wichita Falls, TX
Lawrence	Lawrence, KS	Greater Williamsport	Williamsport, PA
Lawton	Lawton, OK	Eastern Washington	Yakima, WA
Lincoln	Lincoln, NE	York-Gettysburg	York-Hanover, PA
		Youngstown	Youngstown-Warren-Boardman, OH-PA

DATASET

The included CSV comprises a single comprehensive panel dataset underlying all quantitative analysis. The spreadsheet contains all independent, dependent, and normalization variables for 2010 to 2019, inclusive, across the 167 non-overlapping metropolitan statistical areas in our study, each corresponding to a distinct TNC coverage region. These two sets of geographic units were aligned and manually matched with one another according to which pairs most nearly shared centroids and perimeters during the decade under investigation. Spatial matches between most metropolitan areas and their TNC region counterparts are largely self-explanatory based on common names or descriptions. For clarity, all are listed side-by-side, using 2021 service area names, in **Table 1**.

PUBLIC STATEMENTS BY UBER AND LYFT LEADERSHIP

We summarize public statements made by Uber and Lyft leadership first on claims about overall employment, wages and GDP and second on claims about employment and wages in unstable jobs or industries.

Uber and Lyft claims about overall employment, wages, and GDP

The following quotes, cited in chronological order, inform three of our five research questions and hypotheses. *What were the overall economic effects of Uber and Lyft entry in United States cities, specifically on the outcome variables of citywide employment, wages, and metropolitan area GDP?*

“In California alone, Lyft injected $150 million into the economy through people...spending money locally...and creating jobs."
- Logan Green, Chief Executive Officer, Lyft. Remarks at Startup Grind conference, Madrid, Feb 13, 2015.

“Let’s talk about the U.S. specifically. There are not enough people who can reach the income they desire. Wage growth has been fairly anemic. [Uber] is helping push some people into the middle class and provide some security.”
- David Plouffe, Chief Advisor and Board Member, Uber. Remarks at United Nations Job Summit, San Francisco, June 29, 2015.

“Uber allows you to go to job interviews, work on skills, and build your network... now, that’s in the U.S. economy, where the primary issue we face today is underemployment and wage stagnation... The positive economic benefits are not just on [the] driver side. Ridesharing also has a powerful effect on cities, their economies and the people who live in them. Take small businesses, for example. With Uber, you no longer need to be in a particular neighborhood or on a particular street to get customers, and potential customers don’t need a car or a taxi.”
- David Plouffe, Chief Advisor and Board Member, Uber. Remarks at *1776* tech incubator conference, Washington, DC, November 3, 2015

"Lyft has become a powerful driver of economic growth in the Phoenix area by creating flexible economic opportunities for drivers, improved transportation access for passengers, and encouraging local spending.”
- Drena Kusari, Phoenix General Manager, Lyft. Statement to Phoenix Business Journal, December 12, 2016.

“Our riders… represent all adult age groups and backgrounds and use Lyft to commute to and from work… spend more time at local businesses and stay out longer knowing they can get a reliable ride home… As a result of improved freedom to get around, Lyft riders help stimulate local economic activity.”
- Form S-1 Registration Statement, Filed Lyft IPO document. Securities and Exchange Commission,Washington, DC, March 1, 2019.

“When you take an Uber ride, the vast majority of the funds...actually stay in the city and usually go from someone who can afford a ride to someone who needs to earn a living so it’s actually...a pretty strong...driver in terms of money flows in local markets and local cities.”
- Dara Khosrowshahi, Chief Executive Officer, Uber. Remarks to Economic Club of New York, December 4, 2019.

Uber and Lyft claims about employment and wages in unstable jobs or industries
This second collection of quotes, also cited in chronological order, features several of the same Uber and Lyft leaders who spoke to overall economic benefits of TNCS in the previous list. In particular, these statements gesture at a particular improvement for workers employed in less stable jobs, whether temporary, seasonal, transitional, or otherwise benefiting from the flexibility of driving for TNCs. This informs our second question and hypothesis.

“Flexibility is the new stability, and most people aren’t picking one career, one employer… More people are going after their dreams pursuing, you know, careers as, you know, artists, as musicians, as entrepreneurs. Lyft provides a phenomenal platform for people whether they’re in between jobs or pursuing their passion.”
- Logan Green, Chief Executive Officer, Lyft. Remarks at TechCrunch Disrupt conference, New York, May 30, 2015.

"In the Uber world you can use your own car, you make more dollars per hour, and it’s flexible...you don’t have a shift.”
- Travis Kalanick, Chief Executive Officer, Uber. The Late Show with Stephen Colbert, New York, Sep 11, 2015.

“Uber is there to make ends meet for those folks who are maybe finding a transition in their life.”
- Travis Kalanick, Chief Executive Officer, Uber. Squawk Box CNBC interview New York, April 27, 2016.

“Pushing a button, starting work, pushing a button and stopping work. That flexibility in work is, I think, the real breakthrough.”
- Travis Kalanick, Chief Executive Officer, Uber. The Charlie Rose Show, New York, Sep 11, 2017.

“The number one reason that our drivers tell us they love driving for us is because they’re their own boss.”
- Dara Khosrowshahi, Chief Executive Officer, Uber. New York Times DealBook lecture, New York, November 9, 2017.

"Seventy-five percent of the drivers say flexibility is important to them. Most of our drivers are using Uber to fill in the gaps with their other income."
- Olivia van Nieuwenhuizen, Data Scientist, Uber. Statement in Orange County Register, July 26, 2018.

"It's a decent living and, more importantly, it's a very flexible living. You'd be able to drive whatever hours you want."
- Dara Khosrowshahi, Chief Executive Officer, Uber. Economic Club of Washington DC lecture, June 11, 2019.

"The vast majority of our Uber drivers are actually part-time drivers. Uber is...flexible on earnings ...we have drivers who earn when they want to and, you know, many of them are students. Many of them are retirees. Some of them need a side gig."
- Dara Khosrowshahi, Chief Executive Officer, Uber. Remarks to Economic Club of New York, December 4, 2019

"There's going to be much more flexibility both in terms of whether you go to work physically but even if you're going to work there's going to be a lot more flexibility about the hours... and that... creates a more healthy marketplace."
- Dara Khosrowshahi, Chief Executive Officer, Uber. Remarks at Skiff Global Forum, New York, September 28, 2021.

"Look at other, maybe comparable, labor markets. If you look at [the] retail industry, the hospitality-leisure industry, active drivers on Lyft...are coming back five times faster than those industries, so I think a big part of that is the great flexibility we offer."
- John Zimmer, President and Co-Founder, Lyft. Yahoo Finance interview, November 3, 2021.

QUALITATIVE RATIONALE FOR ALTERNATIVE (NON-TWFE) ESTIMATORS

As documented in the paper's literature review, the presence of heterogeneous treatment effects can bias estimates from two-way fixed effects regressions when the timing of treatment is staggered. This necessitates a causal inference method that does not use earlier-treated units as controls for later-treated units. In our context, where all cities are ultimately treated during the study period, we use three alternative modified difference-in-difference estimators, all of which support comparing treated and not-yet-treated units. However, it is first necessary to establish the inappropriateness of two-way fixed effects by critically considering the nature of the treatment variables, namely, the arrival of Uber and Lyft in urban areas.

The staggered timing of treatment is self-evident, as both Uber and Lyft entered different cities in different years. The heterogeneity of treatment across metropolitan areas can be strongly inferred based on economic priors and a broad intuitive understanding of how transportation and land use influence economic outcome variables. Given a static or established set of available transport options, cities maintain an initial equilibrium mode share [10]. With a new option introduced, mode split ultimately (but not immediately) readjusts based on attributes and utility of all alternatives. Attendant to mode share are the social and economic systems tied to transport patterns as well as other time-varying city-specific characteristics (e.g. size, density, proportion of mixed-use

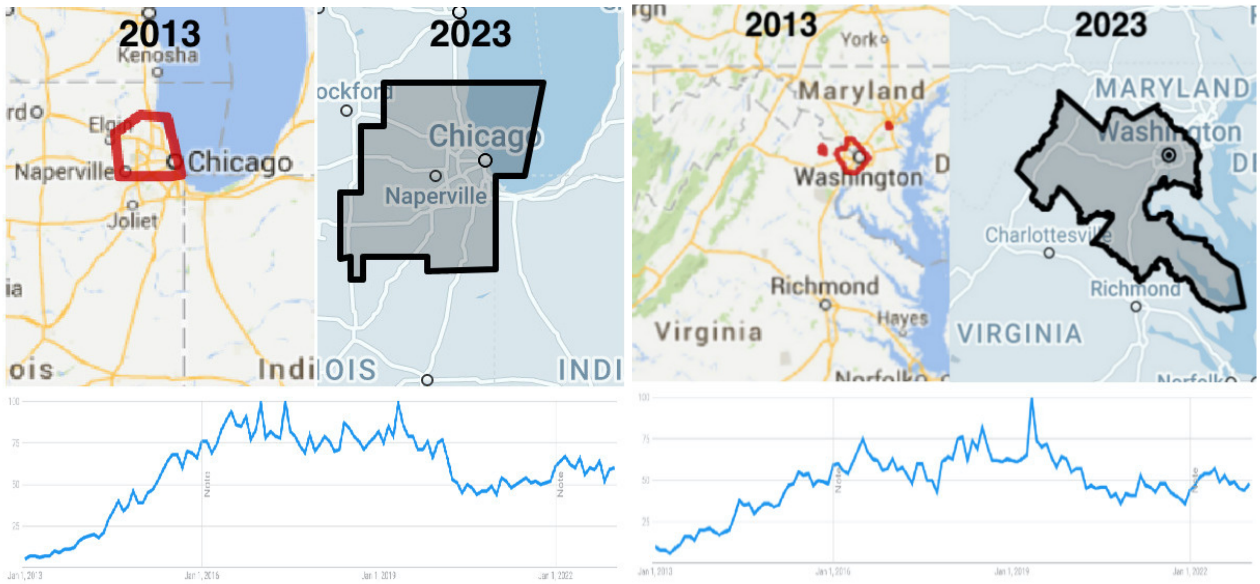


FIGURE 1: Uber Service Area Screenshots and "Uber" Google Search History by Region, Chicago (left) and Washington DC (right), 2013-2023. *Sources: Uber and Google.*

residential and commercial neighborhoods) that may not be captured in city or year fixed effects [7]. In the Callaway Sant'Anna framework, the analogous group- or time-average effects are robust to heterogeneity in treatment effects.

Also threatening to two-way fixed effects is the presence of dynamic treatment across units. Of particular interest here are dynamics that influence the rate of uptake, or the post-treatment "ramping up" of Uber and Lyft in the 2010s. There is strong evidence that the growth of TNCs varied at different rates and at different timescales across metropolitan areas, as illustrated with two representative cities in **Figure 1**.

The side-by-side comparison in **Figure 1** of the Washington, DC and Chicago metropolitan areas, two regions whose initial UberX launches fell earlier in the treatment period, reveals the asymmetric and asynchronous expansion that may not be adequately captured in two-way fixed effects estimation. The cached screenshots on the left, accessed using a web archive of the Uber launch website for each city, represents the coverage area at or immediately following entry. The screenshots on the right, at the same geographic scale, display the coverage areas today. In Chicago, initial Uber coverage favored the city center and the northwestern suburbs, and expanded in all directions. In Washington, DC the initial service area included the District of Columbia, subsets of Arlington and Montgomery Counties, and discrete exclaves at Dulles and BWI airports. Today, it extends nearly 100 miles in multiple directions. The time series underneath each pair of screenshots reflects relative Google search history of the word "Uber" during the same 2013-2023 time period, normalized by city to a 0-100 scale, and reveals a slightly earlier and more sustained search popularity in the Chicago area than in the Washington, DC area. The combined evolution of the spatial service area and search engine input history for Uber in these two cities alone, extrapolated to all regions, constitutes powerful evidence of dynamic treatment effects, and justification for using alternative estimators.

GOODMAN-BACON DIAGNOSTIC

Goodman-Bacon (2021) provides an accompanying quantitative test of whether two-way fixed effects (TWFE) estimation would be problematic [11]. Because the mechanism of bias is incorrect substitution of earlier-treated units as controls for later-treated units, the Goodman-Bacon test detects the degree of effect heterogeneity by metropolitan area, what proportion of units would be incorrectly compared, and, based on the staggered timing of treatment, whether certain units would receive higher weights in the traditional two-way fixed effects estimator. **Figure 2** is a scatterplot showing representative Goodman-Bacon diagnostic results for all outcome variables across our 167 cities, illustrating heterogeneity of sign, magnitude, and weight. Each of numerous triangular data points represent incorrect comparisons of later- to earlier-treated units, and triangular data points further to the right in **Figure 2** reflect constituent 2x2 comparisons weighted more heavily than those closer to the vertical axis. The outcome of the Goodman-Bacon diagnostic reinforces the need to select one or more alternative estimators to TWFE that is robust to the presence of heterogeneity in treatment effects across cities or over time treated. However, for comparison, we still include TWFE event studies in **Figure 6**, recognizing their potential bias.

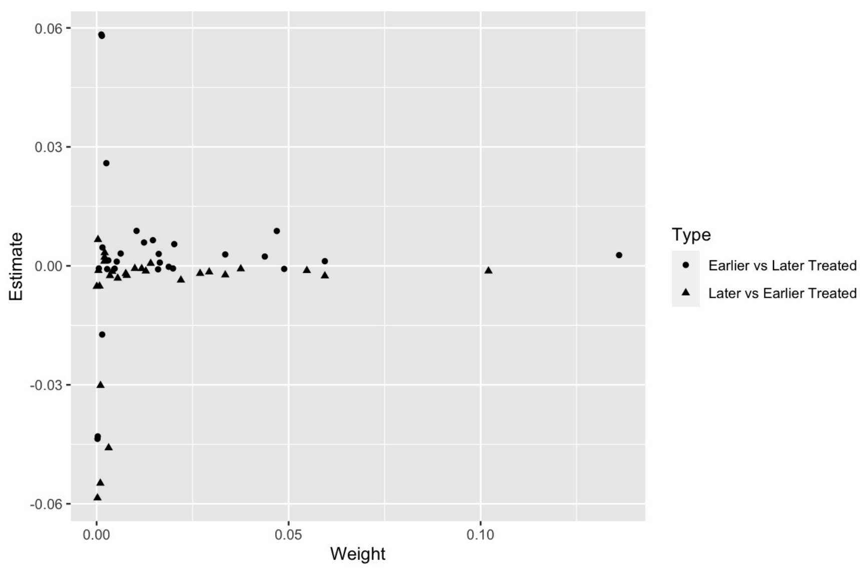


FIGURE 2: Average Goodman-Bacon Diagnostic Results, All Variables

AGGREGATING GROUP-TIME AVERAGE TREATMENT EFFECTS (COHORT, DYNAMIC)

The fundamental Callaway and Sant'Anna estimator (presented in the main paper as **Equation 2** and replicated below) first generates a set of *group-time average treatment effects on the treated*, denoted $ATT(g,t)$, for each outcome variable Y based on cohort g, and calendar year t.

$$ATT(g,t) = E[Y_{i,t} - Y_{i,g-1} | G_i = g] - E[Y_{i,t} - Y_{i,g-1} | D_{i,t} = 0, G_i \neq g] \qquad \forall t \geq g \tag{1}$$

This initial Callaway and Sant'Anna difference-in-differences output is therefore an array of all group-time average treatment effects on the treated in question, each resembling a miniature event study for a single TNC entry year y and subset of cities g treated in that year. For each launch year cohort and calendar year, an ATT is calculated. In **Equation 2**, $ATT(g,t)$ is the group-time average treatment where, for all $t \geq g$, Y is the outcome variable, t is the calendar year, g is the independent variable, the treatment year, G_i is the cohort year in which city i was treated, and $D_{i,t}$ is an indicator variable denoting whether city i was treated at or before calendar year t.

$$ATT(g,t) = E[Y_{i,t} - Y_{i,g-1}|G_i = g] - E[Y_{i,t} - Y_{i,g-1}|D_{i,t} = 0, G_i \neq g] \qquad \forall t \geq g \tag{2}$$

Synthesizing the collection of group-time ATTs into a single interpretable treatment effect, across all cities, requires applying one or more aggregation schemes. The results reported in the paper rely on two aggregation techniques directly endorsed in the Callaway and Sant'Anna paper and accompanying methodological tutorials [6]. Our overall ATTs reflect aggregation by group, calculated using Equation 3, where θ_S^O is the overall effect size across all cities in our dataset, there are a total of T time periods, g and G are the treatment and cohort years, respectively, and $\theta_S(g)$ is the ATT for each cohort.

$$\theta_S^O = \sum_{g=2}^{T} \theta_S(g) P(G = g) \tag{3}$$

The group aggregation θ_S^O expressed in **Equation 3** is, according to Callaway and Sant'Anna, most closely analogous to the overall ATT computed in traditional difference-in-differences analyses [5]. Generating event study analogues from group-time average treatment effects requires a different aggregation scheme that represents the average effect for each year relative to treatment, in our case UberX or Lyft entry. The treatment effect dynamics $\theta_D(e)$, a function of calendar years e before or after treatment, is given by **Equation 4**, where, again, there are T total calendar years, g is the treatment year, and G is the cohort year. In other words, **Equation 4** calculates the average treatment effect on treated cities given e years of TNC exposure.

$$\theta_D(e) = \sum_{g=2}^{T} \mathbf{1}\{g + e \leq T\} ATT(g, g+e) P(g = G|G + e \leq T) \tag{4}$$

The five event study plots in **Figure 3** of the main paper use a universal base period for reference, in which the pre-treatment year $t = -1$ is set to zero and treatment effects interpreted accordingly. While most analogous to TWFE estimation, the Callaway Sant'Anna method also allows for a varying base period. This fixes the reference year to $t = -1$ for treatment and post-treatment years, but in the pre-treatment period uses a rolling base simply of the previous year. We include this varying base period event study as **Figure 3** as an event study in this SI.

ADDITIONAL ESTIMATORS (SUN AND ABRAHAM, STACKED REGRESSION)

The Sun and Abraham estimator effectively combines Goodman-Bacon decomposition with the Callaway and Sant'Anna approach [8]. Just as Callaway and Sant'Anna compile and average across groups by treatment year, the Sun and Abraham method aggregates treatment effects by cohort e. In **Equation 5**, $D^l_{i,t}$ is an indicator variable that is equal to 1 if year t is l years before initial treatment of unit i, $\mathbf{1}\{E_i = e\}$ is a cohort indicator that is equal to 1 for metro areas i who are in treatment year cohort e, $\delta_{e,l}$ is a coefficient representing the average treatment effect on the treated, and γ_i and τ_t are unit and calendar year fixed effects [2].

$$Y_{i,t} = \gamma_i + \tau_t + \sum_e \sum_{l \neq -1} \delta_{e,l}(\mathbf{1}\{E_i = e\}) \cdot D^l_{i,t} + \varepsilon_{i,t} \tag{5}$$

Our third and final estimator relies on stacked regression, drawing on the methods of Cengiz *et al.* (2019) and implemented manually in R with a script modified from Nguyen (2022) and using pre- and post-treatment window of ± 4 years [13]. The primary source of bias in TWFE (**Equation 1** in the main paper) is unwanted constituent 2×2 DiD comparisons that use earlier-treated units as controls for later-treated ones, and are incorrectly aggregated into the overall treatment effect estimate [11]. Subsetting and stacking datasets is a form of gatekeeping; the method applies more restrictive criteria for comparison groups and enables TWFE estimation limited to "clean" controls. However, since treatment effects are identified with ordinary least squares, some OLS biases spotted by Baker (2021), while mitigated, may persist in ATTs estimated by stacked regression [1].

These two additional methods also use not-yet-treated units as controls because all units receive Uber and Lyft service during the ten-year event domain. We also deploy the same simultaneous bootstrap method for unstable employment and unstable earnings, because both are subsets of total employment and earnings, respectively. Again, we use 5,000 replications and a paired t-test to confirm these subsets are statistically significantly different from their superset variables [9].

In **Table 2** we report the point estimates and 95% confidence intervals for all five variables and all three alternative difference-in-differences methods cited in Baker (2021) [1], for the entire event window rather than the [-3, +2] truncation reported in the main paper. These ranges, and the particularly high point estimates, must be interpreted with caution given the inclusion of a small number of much larger cities, treated early and containing data at four, five, and six years following TNC entry. Unlike the main results, these estimates, as illustrated in the event studies on the following pages, effectively draw from a panel that is highly unbalanced in event time.

Given the bias introduced by the known unbalanced panel behind **Table 2** results, we cannot reliably interpret these as effect sizes, nor can we necessarily extrapolate from the narrower event window to a wider one. The magnitudes and uncertainties in the main paper represent a more appropriately conservative subset of the years and cities. However, we note that all three estimators in **Table 2** use markedly different techniques yet return very similar signs, magnitudes, and statistical significances to one another, for all variables in this unbalanced event time domain.

EVENT STUDIES FOR ADDITIONAL ESTIMATORS

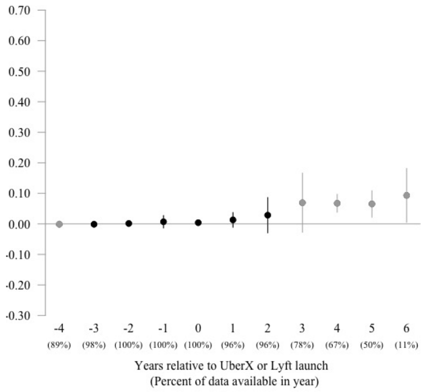
As described in its caption, the event study plots (**Figures 3 - 5**) in the main paper reflect dynamic group-time average treatment effects, aggregated using the Callaway and Sant'Anna method documented in **Equation 4** with a universal base period of $t = -1$. **Figure 3** instead uses the prior year as a reference period before treatment, and year $t = -1$ as a fixed reference year after treatment. These display similar enough effect sizes and error bars to indicate robustness to base year.

TABLE 2: Average Treatment Effects on the Treated, Full Event Time Window

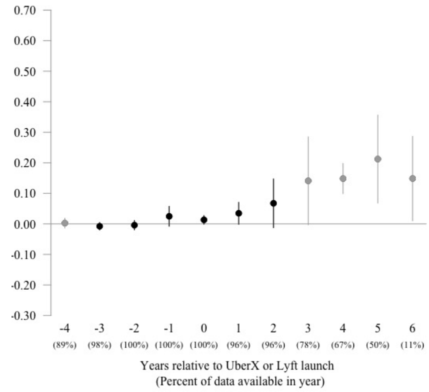
Outcome Variable	**Modified DiD Method**	**Point Estimate**	**95% CI (Low)**	**95% CI (High)**
Log Total Employment per Working Age Population	Callaway & Sant'Anna	0.0295***	0.0023	0.0567
	Sun & Abraham	0.0340***	0.0232	0.0448
	Stacked Regression	0.0371***	0.0238	0.0504
Log Unstable Employment per Working Age Population	Callaway & Sant'Anna	0.0677***	0.0238	0.1115
	Sun & Abraham	0.0789***	0.0564	0.1014
	Stacked Regression	0.0966***	0.0754	0.1178
Log Total Earnings per Working Age Population	Callaway & Sant'Anna	0.0337	-0.0042	0.0716
	Sun & Abraham	0.0278***	0.0086	0.0470
	Stacked Regression	0.0182	-0.0012	0.0376
Log Unstable Earnings per Working Age Population	Callaway & Sant'Anna	0.0611***	0.0021	0.1201
	Sun & Abraham	0.0615***	0.0052	0.1178
	Stacked Regression	0.0624***	0.0242	0.1006
Log Metropolitan or Micropolitan Area GDP per Capita	Callaway & Sant'Anna	0.0536***	0.0211	0.0861
	Sun & Abraham	0.0678***	0.0239	0.1117
	Stacked Regression	0.0607***	0.0405	0.0809

The event studies in **Figures 4** and **5** correspond to the other two estimators whose ATTs are reported in **Table 2**: Sun & Abraham and stacked regression, respectively. In addition, we provide a set of event studies, using two-way fixed effects, for all variables in **Figure 6**, which we report alongside the major caveats noted earlier in this SI and in the main paper methods section: in particular, staggered treatment timings with heterogeneous and dynamic effects.

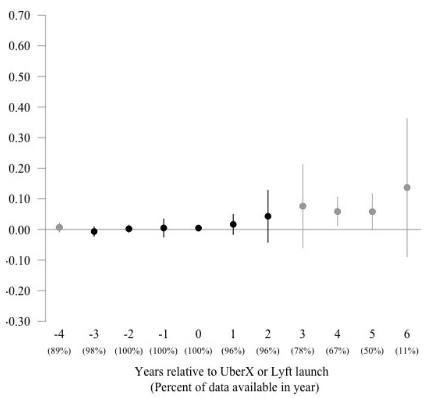
Just as a broadly similar pattern of magnitude and statistical significance was evident in ATTs for each outcome variable across all estimators (Table 1 in the paper), the same is largely true in the Sun & Abraham and stacked regression event studies, with no statistical significance during or before the launch year, and positive and a statistically significant treatment effect measured for all five variables by the fourth year after UberX or Lyft entry. The years in gray, like in the main paper event studies, should be interpreted with greater caution and with particular attention to the percentage of metropolitan areas represented in that year relative to treatment.



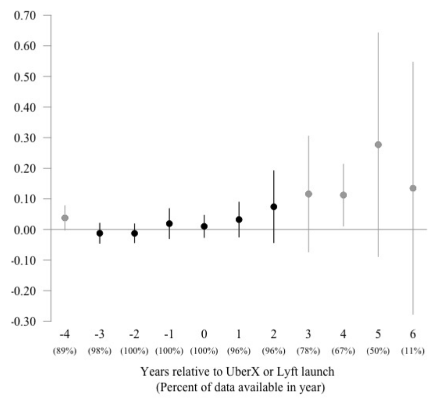
(a) Total Employment



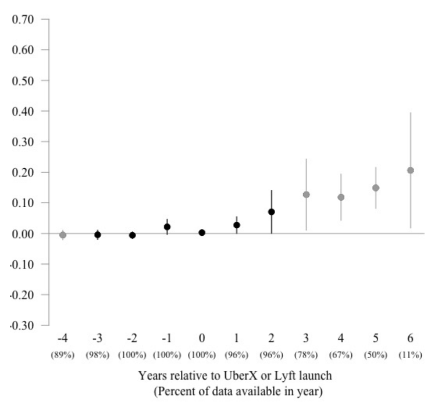
(b) Unstable Employment



(c) Total Earnings

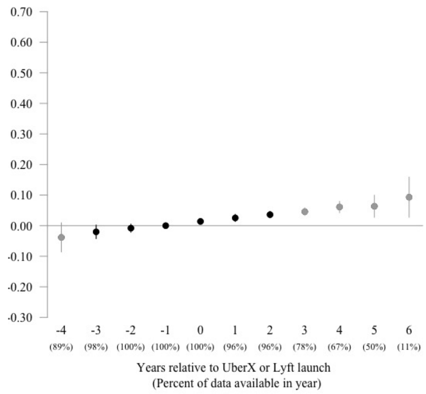


(d) Unstable Earnings

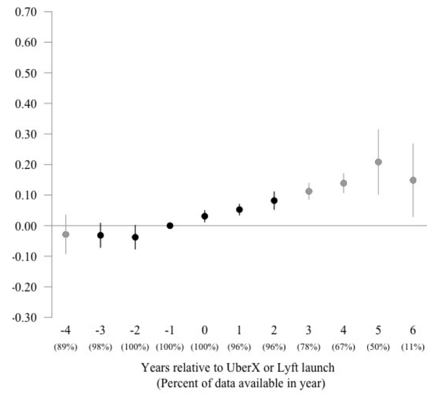


(e) GDP

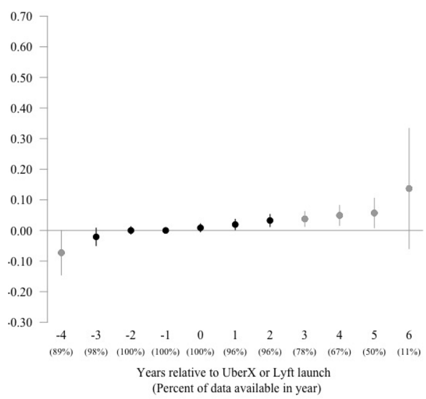
FIGURE 3: Callaway & Sant'Anna *varying base period* event study results for average treatment effects, by year, for (a) total employment, (b) unstable employment, (c) total earnings, (d) unstable earnings, and (e) GDP. All variables are normalized by population (total or working-age) and log transformed. Only pre-treatment estimates and errors depend on base year. All post-treatment point estimates and confidence intervals are identical to those in **Figure 3** of the main paper. Estimates are shown in gray for years with limited data availability outside the [-3, +2] event time window.



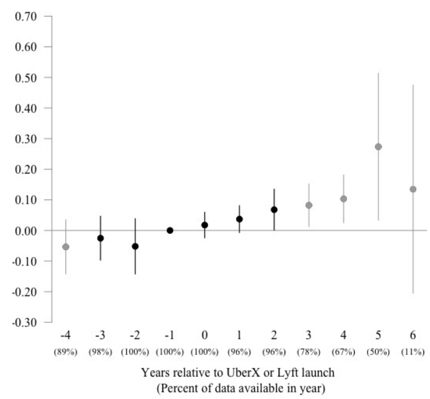
(a) Total Employment



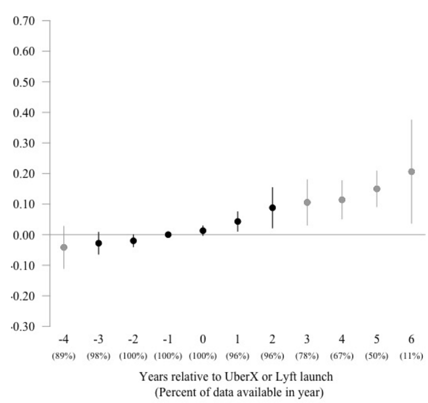
(b) Unstable Employment



(c) Total Earnings

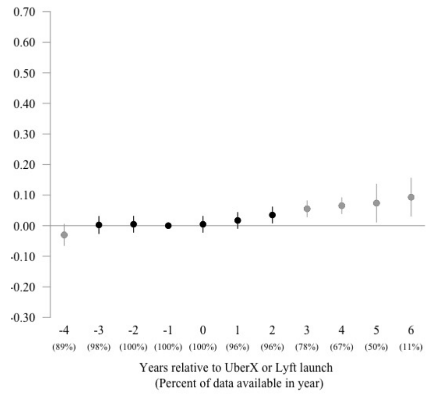


(d) Unstable Earnings

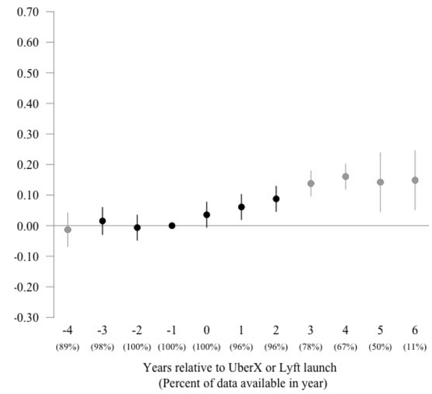


(e) GDP

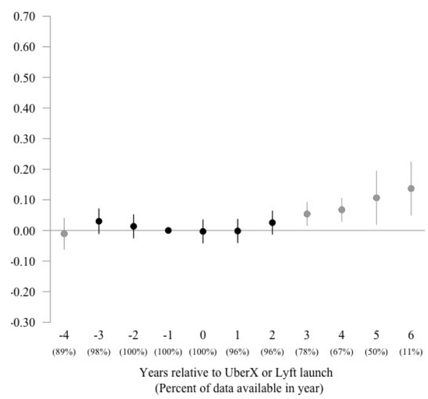
FIGURE 4: Sun & Abraham event study results for average treatment effects, by year, for (a) total employment, (b) unstable employment, (c) total earnings, (d) unstable earnings, and (e) GDP. All variables are normalized by population (total or working-age) and log transformed. Estimates are shown in gray for years with limited data availability outside the [-3, +2] event time window.



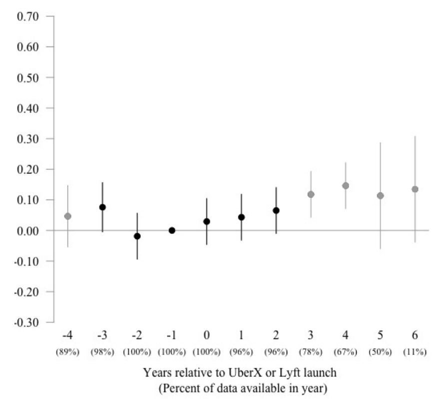
(a) Total Employment



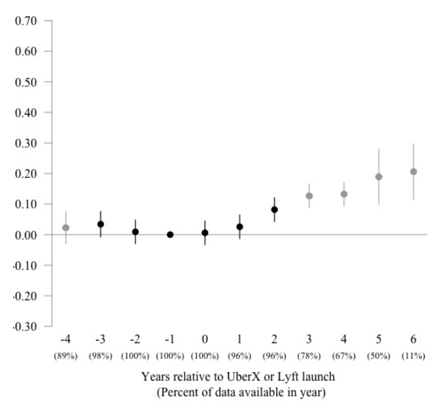
(b) Unstable Employment



(c) Total Earnings

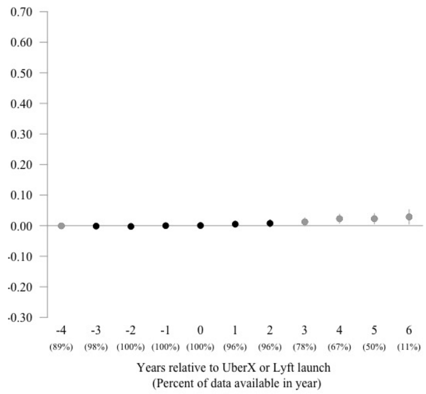


(d) Unstable Earnings

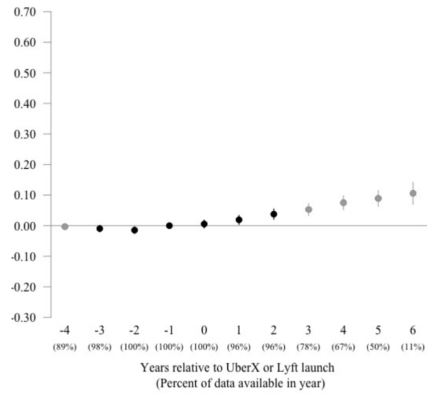


(e) GDP

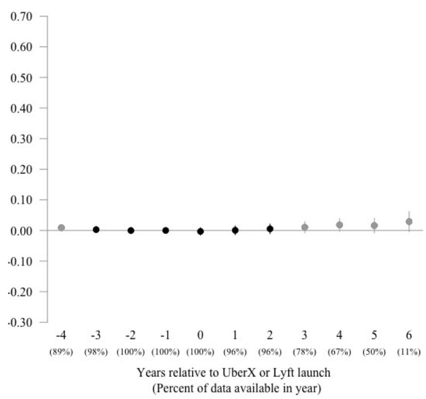
FIGURE 5: Stacked regression event study results for average treatment effects, by year, for (a) total employment, (b) unstable employment, (c) total earnings, (d) unstable earnings, and (e) GDP. All variables are normalized by population (total or working-age) and log transformed. Estimates are shown in gray for years with limited data availability outside the [-3, +2] event time window.



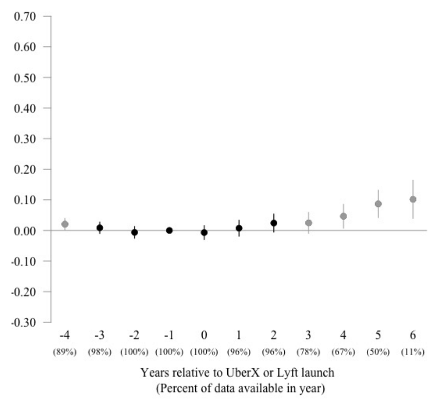
(a) Total Employment



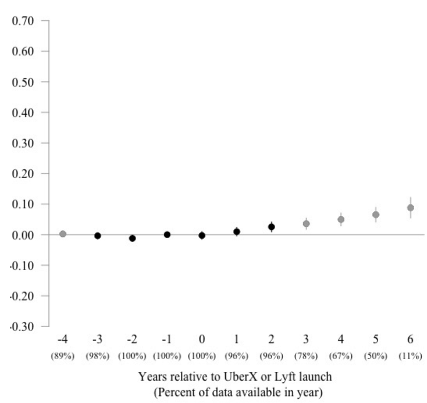
(b) Unstable Employment



(c) Total Earnings



(d) Unstable Earnings



(e) GDP

FIGURE 6: Two-way fixed effects event study results for average treatment effects, by year, for (a) total employment, (b) unstable employment, (c) total earnings, (d) unstable earnings, and (e) GDP. All variables are normalized by population (total or working-age) and log transformed. Estimates are shown in gray for years with limited data availability outside the [-3, +2] event time window.

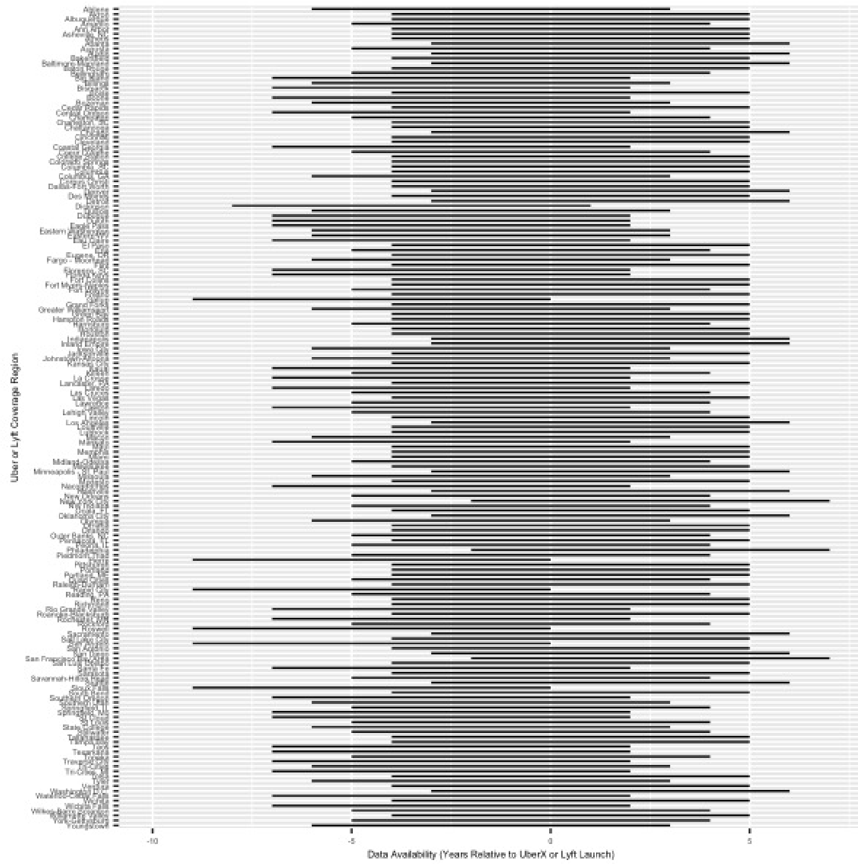


FIGURE 7: Data availability by MSA and year before or after launch

DATA AVAILABILITY AND EXTENDED EVENT STUDIES

All estimates and error bars are in gray outside the previously mentioned [-3, +2] window, based on the data availability illustrated in the main paper histogram (**Figure 2**) and in **Figure 7** above. There are visibly sharp reductions in data availability in the fourth years before and after UberX or Lyft launch, and **Figure 7** shows all years of availability at the individual metropolitan area level.

IDENTIFICATION ASSUMPTIONS

We develop and examine a composite list of assumptions under which the methods that we use produce unbiased causal inference estimates, using Callaway and Sant'Anna (2021) as a template, made explicitly or implicitly across the three seminal papers for each respective alternative estimator. We then consider these assumptions in context of the 2010-2019 TNC dataset and comment on how satisfactorily our three models ostensibly meet the necessary prerequisites. As certain assumptions appear more sound than others for our underlying data, this section cumulatively introduces caveats to interpreting our final results. Beyond the assumptions identified in Callaway and Sant'Anna (2021), we discuss one additional possible challenge to identification.

Callaway & Sant'Anna Assumptions

Irreversibility of treatment

The assumption of irreversibility of treatment requires that once a unit is treated it stays treated in all future time periods. This holds in our data with one exception: Uber and Lyft were briefly banned in Austin, TX from May 2016 to May 2017. During this time an alternative TNC service, RideAustin, continued to operate. The time period of our analysis is the calendar year, and we observe no calendar years without treatment.

Random sampling

$\{Y_{i,1}, Y_{i,2}, ..., Y_{i,\tau}, X_i, D_{i,1}, D_{i,2}, ..., D_{i,\tau}\}_{i=1}^{n}$ is independently and identically distributed, where τ is the last time period, $Y_{i,t}$ are the potential outcomes, and $D_{i,t}$ are the treatment indicators. Intuitively, this assumption states that the sampling of units is not a function of the path of treatment status. In our case, we did not dictate which cities entered our sample on the basis of whether or when Uber and Lyft entered the city (i.e., we include all cities for which we have data on the relevant outcomes). This assumption also implies that the treatment status of one unit does not affect the outcome of another unit (no spillover). While availability of TNCs in other cities could plausibly affect travel between cities, we do not expect treatment of one city to meaningfully affect economic outcomes in other cities.

Limited treatment anticipation

The assumption of limited treatment anticipation requires that treated groups do not anticipate and respond to treatment more than *n* years in advance. The Sun & Abraham method imposes a stricter assumption of no anticipation. Given the limited service regions during early launch and the speed and secrecy of launch decisions, we would not expect regional markets to anticipate TNC entry in a way that would affect the indicators that we measure, and the absence of statistically significant pre-treatment effects for any estimator provides evidence consistent with this assumption.

Conditional parallel trends

The assumption of conditional parallel trends based on not-yet treated groups requires that, after conditioning on any covariates, the differences between treated and control groups in pre-treatment periods hold constant over time. Our event studies reveal no statistically significant pre-treatment effects for any variables or methods, which is consistent with parallel trends.

Overlap

The assumption of overlap requires that the propensity of each unit to be treated is bounded away from 0 and 1, meaning there are no systematic processes that would guarantee or prohibit a unit from being treated. We are not aware of any such systematic factors. Although policy restricted operation of Uber and Lyft in Austin, TX from May 2016 to May 2017, the policy itself changed over time, and all units had the potential to be treated or not treated.

Correlated entry of other gig economy platforms

Finally, we must consider that TNC entry timing could be correlated with entry of other smartphone-based gig economy platforms, such as Airbnb, Grubhub, TaskRabbit, or other electronically-mediated services. While there are varying definitions of the scope of the gig economy and how to measure a particular industry's market share, TNCs were estimated in 2018 to comprise roughly half of app-based economy activity and revenue in the United States [12]. Importantly, like most United States labor data, the QWI variables may not account for Uber and Lyft drivers consistently across cities or comprehensively within them. The core indicators used (beginning-of-quarter and stable employment and earnings) exclude certain highly unstable jobs and individuals who are not employed at either boundary between quarters.

If TNCs systematically launched in similar regions at similar times as other key players in the peer-to-peer app-facilitated economy, then our estimates may capture effects attributable to a combination of TNCs and other concurrent entry effects [3] (If TNCs launched at different times than other key players, then treatment effects of these other players could potentially violate the parallel trends assumption, though our results are all consistent with the parallel trends assumption). Compared to the more readily available spatial and temporal resolution of TNC entry data, we could not find systematic nationwide data on launch years or market penetration for these other services to assess the degree of correlation. Beyond the confounding risks from simultaneous entry and growth, gig platforms may further interact with and complement one another. In Austin, for example, joint Uber and Airbnb economic effects were estimated to be greater than the sum of each service alone [4].

References

[1] A Baker, D Larcker, and C Wang. *How much should we trust staggered difference-in-differences estimates. Harvard Business School*. 2021.

[2] Andrew C Baker, David F Larcker, and Charles CY Wang. "How much should we trust staggered difference-in-differences estimates?" In: *Journal of Financial Economics* 144.2 (2022), pp. 370–395.

[3] Andrew T Bond. "An app for that: Local governments and the rise of the sharing economy". In: *Notre Dame L. Rev. Online* 90 (2014), p. 77.

[4] Gordon Burtch, Seth Carnahan, and Brad N Greenwood. "Can you gig it? An empirical examination of the gig economy and entrepreneurial activity". In: *Management Science* 64.12 (2018), pp. 5497–5520.

[5] B Callaway and PHC Sant'Anna. *Introduction to DiD with Multiple Time Periods*. 2022.

[6] Brantly Callaway and P Sant'Anna. *Getting started with the did package*. 2021.

[7] Robert Cervero. "Built environments and mode choice: toward a normative framework". In: *Transportation Research Part D: Transport and Environment* 7.4 (2002), pp. 265–284.

[8] Scott Cunningham. Presentation by Scott Cunningham at CodeChella 2021 on Sun and Abraham estimator. Aug. 2021.

[9] Daniel J Eck. "Bootstrapping for multivariate linear regression models". In: *Statistics & Probability Letters* 134 (2018), pp. 141–149.

[10] Feras El Zarwi, Akshay Vij, and Joan L Walker. "A discrete choice framework for modeling and forecasting the adoption and diffusion of new transportation services". In: *Transportation Research Part C: Emerging Technologies* 79 (2017), pp. 207–223.

[11] Andrew Goodman-Bacon. "Difference-in-differences with variation in treatment timing". In: *Journal of Econometrics* 225.2 (2021), pp. 254–277.

[12] Lawrence Mishel. *Uber and the labor market: Uber drivers' compensation, wages, and the scale of Uber and the gig economy*. 2018.

[13] Mike Nguyen. *A Guide on Data Analysis*. Bookdown, 2022.