Transportation
https://doi.org/10.1007/s11116-023-10425-w

What stay-at-home orders reveal about dependence on transportation network companies

Lily Hanig[1] · Destenie Nock[1,2] · Corey D. Harper[2,3]

Accepted: 3 September 2023

Abstract

Transportation Network Companies (TNC) such as Uber and Lyft set out to provide transportation not fulfilled by private vehicles or public transit. The social value of TNCs for essential trips (i.e., necessary trips that cannot be fulfilled by another mode of transportation) is difficult to discern in normal conditions. The COVID-19 stay-at-home order is used as a natural experiment to investigate the heterogeneous ability to avoid TNCs by income areas of trip origins. We measure the sensitivity of different populations' ability to respond to policies and to avoid TNC trips (e.g., early stay-at-home orders) using a difference-in-difference style regression. Previous studies have indicated that under normal conditions TNCs primarily serve high-income areas, indicating that TNCs may not be improving transportation equity but instead serve as an additional mode of transportation for passengers with multiple options. We fill a gap in the literature by evaluating the role TNCs play in serving unavoidable and essential trips. We find that high-income community areas showed greater sensitivity to the stay-at-home order with a 56% greater decrease in TNC ridership during the stay-at-home order compared to low-income community areas. Specifically, TNC trips from high-income areas decreased by 80%. This indicates that although riders from high-income community areas might make up the majority of trips in normal conditions, low-income community areas are less able to adapt to stay-at-home orders because of a higher degree of non-flexible and essential jobs or less access to TNC alternatives like private vehicles and public transit.

Keywords Transportation network company · Uber · Equity · Transportation equity · COVID-19

✉ Lily Hanig
lhanig@andrew.cmu.edu

1 Engineering and Public Policy, Carnegie Mellon University, Pittsburgh, USA

2 Civil and Environmental Engineering, Carnegie Mellon University, Pittsburgh, USA

3 Heinz School of Information Systems and Public Policy, Carnegie Mellon University, Pittsburgh, USA

Published online: 07 October 2023

Introduction

The remote work economy spurred by the COVID-19 stay-at-home orders has drastically changed travel patterns across the US; public transit has been slow to recover even years out from the onset of the pandemic (Zipper 2021). During the pandemic's initial lockdown period in the US (mid-March to late-April) ridership across all forms of travel (e.g., buses, TNCs) drastically reduced. Public transit ridership dropped by as much as 90% in some urban areas during initial stay-at-home orders (Liu et al. 2020; Qi et al. 2021), streets were empty, and schools shut down. Transportation network company (TNC) ridership was no exception. Many TNC riders, particularly high-income riders, switched to alternative modes (e.g., a private car) and began working from home if commuting was non-essential (Liu et al. 2020; Qi et al. 2021). The dramatic reduction in ridership creates a natural experiment that can illuminate essential TNC trips that are more difficult to discern during non-pandemic conditions.

During the first few months of the COVID-19 pandemic when stay-at-home orders were in-place, most trips were essential (City of Chicago 2022; Brodeur et al. 2021). Here, essential trips are defined as work, home, healthcare appointments, pharmacies, or the grocery store (City of Chicago 2022). Given the city-mandated stay-at-home order and concerns about infection during the first six weeks of the stay-at-home order from March 18th to April 29th, if people had access to private vehicles they used them instead of taking TNCs or other public options (Ozbilen et al. 2021). Indicating, if a person took a TNC trip during the early stages of the pandemic, they likely lacked access to private transportation. The early period of the COVID-19 pandemic provides a unique opportunity to assess essential TNC ridership as non-essential trips were kept to a minimum. Our research questions are as follows: (1) what is the magnitude of essential TNC trips that still occur during the stay-at-home order compared to pre-pandemic trips? (2) Is there a heterogeneous response in TNC ridership to the stay-at-home orders by income? and (3) have time-of-day trip patterns changed during the stay-at-home order?

Some previous studies show that TNC trips have a higher positive correlation with high-income areas than with low-income areas and areas that have limited transportation options (Grahn et al. 2020; Barajas and Brown 2021) indicating TNCs primarily act as a service of convenience for high-income riders who are more likely to have access to other modes of transportation (Grahn et al. 2020; Barajas and Brown 2021). To understand what role TNCs play in transportation equity, several papers investigate the relationship between TNCs and transit. Ward et al. (2019) finds a slight correlation between TNC entry and decreased vehicle registrations, indicating there could be some positive impacts on emissions and transit access with TNC entry (Ward et al. 2019). Some papers show that the entry of TNCs across cities has a net increase in public transit ridership, potentially by providing riders with a backup option in case of disruptions in public transit service, which makes them more assured in their decision to take public transit over using a private vehicle. (Hall et al. 2018; Nelson and Sadowsky 2019). Although high-income and choice riders with multiple transportation options may be the bulk of TNC riders in normal conditions, there is a gap in the literature assessing the dependency on TNC trips for essential purposes by income groups. Despite much of the literature indicating that the majority of riders do not depend on TNCs to get to their destination, there may be a minority of riders that are low-income or lack alternatives to TNCs to make their trip.

Under normal conditions, it is difficult to discern TNC-dependent riders from choice or non-essential riders; TNC trips do not get designated by the rider as essential or non-essential and the commuter patterns of essential workers are less likely to follow the 9-to-5 pattern that make up the bulk of pre-pandemic commuter transit trips (Transit Center 2020). Additionally, varying types of trips could be categorized as essential beyond trips to work (e.g., medical appointments and grocery shopping). The stay-at-home order provides a natural experiment to assess the level of essential and non-essential trips occurring via TNCs. We use the initial stay-at-home order in Chicago in March of 2020 during the COVID-19 pandemic as the policy shock to assess the difference in ridership across community areas by income from before and after the stay-at-home order. We measure the heterogeneous response to the stay-at-home order across income groups to understand what role TNCs play in transportation equity and providing essential service.

Literature review

It can be difficult to discern in non-pandemic conditions which trips are essential and which trips could not be served by an alternative mode. Some previous studies (Barajas and Brown 2021; Ngo et al. 2021) on essential trips classify trips as essential and non-essential by the day and time of the trip; these papers make the assumption that trips on weekdays for the typical morning and evening commuter peak are riders commuting to work (Barajas and Brown 2021; Ngo et al. 2021); however, this may not encompass night-shift commuter trips (e.g., nightshift nurses) or other odd-hour work trips. According to the Transit Center, 41% of US essential workers commute outside of the typical 9 to 5 commuter times (Transit Center 2020), and these workers are more likely to be low-income or an underrepresented minority (TransitCenter 2020; Kantamneni 2020). Papers using typical office worker commute windows as a proxy for essential trips prior to the pandemic leave a gap in the literature for assessing essential trips that occur outside of typical commuter hours. Several papers use the COVID-19 pandemic to contribute to the literature on the essentialness of trips by income (Pawar et al. 2021; Hook et al. 2021; Brough et al. 2021; Iio et al. 2021) and find that travel in high-income areas decreased more than travel in low-income areas. However, there is a gap in the literature specifically assessing the dependence on TNCs by income. We contribute to this body of literature, additionally finding a higher degree of essential travel from low-income areas and specifically filling in a gap in the literature by investigating the dependence of high and low-income travelers on TNCs for essential trips, whereas other studies focus on general mobility patterns or public transportation.

We assume that in the early phase of the stay-at-home order, private modes of transportation (e.g., cars, bikes) are preferred for physical distancing, and if a TNC trip still occurs, the rider did not have a better option (Ozbilen et al. 2021). We also assume that TNC trips occurring during the initial six weeks of the stay-at-home order are essential, even if these trips occur outside of peak weekday commuter travel times (City of Chicago 2022; Brodeur et al. 2021).

TNCs provide flexibility in transportation options, therefore it is difficult to discern which trips could not be completed with an alternative mode. There are debates on the transportation equity benefits and drawbacks of TNCs, we highlight literature on both sides of the argument here. One study states that low-income, minority, and disabled TNC riders, on average, experience longer wait times and more frequent ride cancellations (Jin et al. 2019). The literature is inconclusive on the degree to which TNCs complement public transportation by filling in gaps in service or TNCs compete with public transportation by taking away ridership and impinging on public services that transit-dependent riders rely on. Some studies indicate that TNCs are mostly used by populations who are more likely to have access to other forms of transportation either through a private transportation mode or living in walkable and bikeable areas with high access to transit (i.e., higher income earners and people with higher levels of education) (Grahn et al. 2020; Barajas and Brown 2021). In contrast, several papers find that the introduction of TNCs to city(s) shows a correlation with decreased public transportation use (Diao et al. 2021; Erhardt et al. 2021; Ngo et al. 2021). These papers indicate that TNCs serve passengers who likely have alternative modes of transit they could take instead. The literature is inconclusive, which leaves room for analysis of what role TNCs play in improving, maintaining, or impairing transportation equity. We contribute to the literature by investigating the dependence of high and low-income travelers on TNCs for essential trips.

Li et al. (2019) assesses the degree to which different policies impact drivers and TNC financial viability (Li et al. 2019); specifically, this paper assesses policies for driver protections and finds that TNC drivers take on the highest degree of risk given demand uncertainty (Li et al. 2019). We contribute to the TNC policy-response literature by using the stay-at-home order mandate as an opportunity to assess the response to policy by the rider population, rather than focus on drivers. One paper by (Brown and Williams 2021) uses the lock-down period to compare Uber ridership by low and high-income riders (by trip origin) for essential and non-essential trips. (Brown and Williams 2021) adopts Uber's essential trip designations to define if a trip is essential based on the destination the passenger places into their Uber app. For example, if a rider types 'Hospital' into the app, that trip would be flagged as essential. One limitation in (Brown and Williams 2021) is that there are trips that needed to occur but lacked a specifically designated 'essential' destination, such as the return trip home from the hospital, which would be excluded from their analysis. We fill this gap by defining an 'essential TNC trip' as any trip that could not be avoided and could not be completed with alternative modes during the stay-at-home order. We assume that all TNC trips occurring in the first six weeks of the stay-at-home period during the pandemic are essential and unavoidable. In our analysis we use TNC data collected by the city of Chicago from all TNC providers (e.g., Uber and Lyft).

Case study

Chicago background and the pandemic

We use Chicago, IL as our case study due to the city requiring all TNC companies to make their trip data publicly accessible (City of Chicago 2020a). Chicago has 2.7 million residents (US Census 2021), making it the third largest metropolitan area by population in the US. The Chicago Transit Authority reports an average of 1.6 million transit rides (e.g., bus and rail) each week, before the pandemic. While Chicago has a healthy public transit system there is a strong presence of TNCs. Uber entered Chicago in 2011 (Rao 2011), and other TNCs followed. It is estimated that before the pandemic there were roughly 2 million TNC trips across the city each week (City of Chicago 2020a).

In Chicago in mid-March of 2020, cases of COVID-19 were still very low in the city, but stay-at-home policies had high compliance out of fear of disease spread (City of Chicago 2022; National Center for Immunization and Respiratory Diseases (U.S.). Division of Viral Diseases 2020; City of Chicago 2020b). Non-essential trips, such as visiting friends and family, as well as outings to bars, restaurants, or parties were largely avoided. However, essential trips like the grocery store or healthcare appointments that could not be completed virtually still required public transit and TNC usage in households without a private car. We hypothesize that low-income community areas have TNC ridership that is 'essential' at a higher rate and is less sensitive to policy changes in the city surrounding COVID-19. This may be due to barriers to using online and virtual services (e.g., lack of internet, inability to use food vouchers online), fewer alternative modes, less consistent public transit access, lack of a private vehicle, or lack of affordable or accessible delivery services. Higher dependency on TNCs could also be due to low-income and residents without private cars working as essential workers at higher rates (Transit Center 2020). These essential trips still occur in non-pandemic times but are potentially masked by the large quantity of high-income, choice, and non-essential TNC rides. Thus, the early high-compliance stay-at-home order can reveal the degree of essential trips during non-pandemic conditions, as well.

On March 15th 2020, the Center for Disease Control (CDC) announced guidance discouraging large gatherings of 50 or more people (National Center for Immunization and Respiratory Diseases (U.S.). Division of Viral Diseases 2020). Closely following the CDC guidance, many cities, including Chicago, placed stay-at-home orders in the days that followed to prevent the spread of COVID-19 (March 18th, 2020 for Chicago) (City of Chicago 2022). In the wake of the stay-at-home orders all schools and jobs that could go remote did so. We use the six weeks leading up to the CDC guidance (March 15th) as the before period to the pandemic and the six weeks following the onset of the stay-at-home order (March 18th) as the stay-at-home period. We exclude March 15th to 18th as the transitional period to the onset of policies mandating people to stay at home. The city of Chicago placed a change in the tax structure for TNCs to help mitigate congestion on January 5th, 2020; we use only the six weeks leading up to the stay-at-home order (starting February 1st) to not have a confounding impact from the change in the tax structure for TNCs.

The rate of COVID-19 in the city of Chicago remained very low during the first six weeks of the stay-at-home order (less than 1% of the city) (City of Chicago 2020b). This implies that the decrease in ridership observed can be largely attributed to policy compliance and fear, than due to COVID-19 health complications directly causing passengers to not be able to ride. In summary, our study fills a gap in the literature by investigating the dependence of high and low-income travelers on TNCs, using COVID-19 policies as a natural experiment to study the impact of stay-at-home orders directing people to only take essential trips.

Data

The city of Chicago requires TNC companies like Uber, Lyft, and Via to report data of trip origins and destinations (at the census tract and community area levels), pick-up and drop-off time (rounded to the 15-minute), price (rounded to 50-cents), tax, length of trip in distance and minutes, and if it was a pooled trip (shared between two unacquainted passengers). Chicago has 866 census tracts within 77 community areas. We exclude two community areas in our analysis because they contain airports.

The Chicago TNC data provides trip-level information on origins and destinations at the census tract and community area level, but the data do not contain socioeconomic or demographic information on the rider. As a result, we use the per capita income of the trip's origin community area from the 2018 American Community Survey (CDC/ATSDR

SVI Data and Documentation Download 2021). We use walk score, bike score, and transit score from Walkscore.com matched to community areas for a rating from 1 to 100 of the walkability, bikeability, and transit access of each community area. This allows us to better understand the ease of access to transit, walk, or bike modes for a trip from each area. This provides a spatial proxy of income (See the Conclusion for a discussion of limitations).

We use TNC trips aggregated to the day per community area per capita as the dependent variable. For passenger anonymity, if a trip were the only trip to occur in a 15-min window for a census tract origin or destination, the trip still appears in the database, but the origin or destination census tract is not given (the community area is still given, unless it's the only trip in the community area, too). Trip frequency decreased overall during the stay-at-home order, making trips more likely to be the only one in a census tract in a 15-min window, and therefore more likely to be excluded. To offset the biased data exclusion, we aggregate to the community area level because the community area is less often excluded.

Methodology

We first conduct an exploratory analysis using descriptive statistics on 20% of data to formulate our model; we investigate the change in magnitude of TNC trips before and during the stay-at-home order to identify the heterogeneous response to policy by income, and the revealed time-of-day pattern changes. After that, we run regressions on the remaining 80% of data to assess the difference in TNC trips before and after the stay-at-home order and investigate the heterogeneous response to the stay-at-home order across incomes. Therefore, we look at the difference across time before and after the stay-at-home order and the difference in response across per capita income of trip origins. The interaction term (per capita income of the origin by the stay-at-home order) measures the difference in the difference of responses (e.g., TNC usage). Finally, we do a posthoc analysis looking at the origin-destination pairs of trips by time of day before and after the pandemic to investigate changes in trips by time of day from low to high-income areas versus high to low-income areas before and after the stay-at-home order.

Data processing

The overview of the data aggregation process for each dataset by temporal and spatial resolution is shown in Fig. 1. The city of Chicago gives access to TNC trip data at 15-minute intervals with both the census tract and community area of the origin and the destination. Demographic data are averaged using population-weighting to the community area. The trips are also aggregated to the day for each community area. The result is the total trip count of TNC trips that originate per community area per day with community area mobility and demographic characteristics tied by community area.

Descriptive statistics

The dataset has trips per day from February 1st to April 29th, 2020. Aggregating trip origins over the course of each day for each community area yields N = 6375 trip totals, which includes a total of 13.2 million TNC trips over the three months in the city of Chicago (2.7 million trips in the sampled 20% of data used for model formulation, and 10.5 million trips in the 80% used for testing data). There was an 84% decrease in trips from

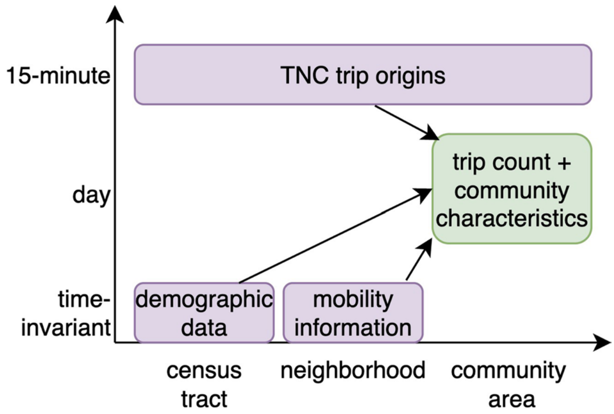

Fig. 1 Diagram of raw and aggregated data sets by time and spatial granularity

February to April. Compared to February, April's time-of-day trends had less variation in origin-destination trends across the day and demonstrated less pronounced commute-hour peaks (see Origin-Destination Pairs for further details on time-of-day ridership). Table 1 shows the mean, standard deviation, 5th percentile, and 95th percentile for trip origins in the 20% of data used for model formulation for the independent variables. Table 1 shows a relatively low mean per capita income ($30,400) and mean percent ethnic minority at the community area level. Additionally, the low average for non-vehicle ownership combined with transit, walk, and bike scores in the 60–75% range suggest a dependence on private vehicles and an overall lack of public options. See Appendix A for further descriptive statistics and correlations. The 75 community areas range in population from 2000 to 110,000 residents.

Formulation

We use a difference-in-difference (DiD) style regression to assess the heterogeneous impacts of the stay-at-home order policy across income groups to investigate changes in ridership before and after the first COVID-19 stay-at-home order in Chicago. We run the regression for income split into low and higher-income areas at the median and as a continuous difference in difference by income (Callaway et al. 2021).

We use 20% of the trips aggregated to the day for model formulation and to look at descriptive statistics to develop a model hypothesis. We then use the remaining 80% of the

Table 1 Descriptive statistics for the dependent variable and the time-invariant independent variables. N = 6,375 for trips aggregated to the day and community area

	Mean (%)	Std. deviation (%)	5%pc	95%pc
Per capita income	30,400	17,800	13,700	72,900
Percent no vehicle	10.4	6.5	2.2%	20%
Percent ethnic minority	73.0	26.7	21%	99%
Transit score	63.5	11.9	42%	82%
Walk score	73.1	12.1	50%	92%
Bike score	70.1	13.1	53%	91%

data to run the regressions and test both models with five-fold cross-validation to verify that the models are not overfitted to the training data. We run two regressions, both using community area-level fixed effects (i.e., per capita income, percent ethnic minority, walk score, bike score, transit score), day-of-the week, the stay-at-home order dummy, and the interaction term between income and the stay-at-home order. In the first formulation, we split the community areas into low-income and higher-income groups set at the median income. This is done to capture the heterogeneous impacts between low and higher-income groups. In the second formulation, we use the income of the area as a continuous variable in the interaction term, to capture the incremental heterogeneous impacts of income on the ability to avoid trips. We detail our hypothesis testing, fitting, and model formulation in Eqs. 5–9 in Appendix B.

The six weeks prior to the CDC guidance from February 1st, 2020 to March 15th, 2020 are the pre-treatment period. March 15th through March 18th are excluded as the transitional period because national and local guidance was encouraging people to stay home but the formal mandate was not in-place yet. The treatment period ($\sigma_t = 1$) in this study is the six weeks following the onset of the Chicago stay-at-home order on March 18th, 2020. The 77 community areas (i) by per capita income are the comparative groups. Both groups get the treatment, but the effect being measured is the difference between low and higher-income (Γ_i) and the incremental difference in sensitivity to the mandate per income as a continuous variable (γ_i), similar to a dose-response study (Callaway et al. 2021). This study uses the stay-at-home order as the mandate that the different treatment groups are receiving and the heterogeneous ability to curb ridership as the measured interaction.

The full regression for both formulations (Eq. 1 and Table 2) are below. The log of trip counts per community area per day (Y_{it}) normalized to the population in the community area is the dependent variable for both the continuous and discrete income group formulation of the regression. The time-variant fixed effects are the day of the week (δ_t) and the stay-at-home order (σ_t). The time-invariant fixed effects are the log transformation of per capita income (γ_i), percent non-vehicle ownership (ν_i), percent ethnic minority (μ_i) walk score (ω_i), and transit score (τ_i) at the community area level. The parameter of interest

Table 2 Overview of variable symbols

Variable	Symbol	Manipulation
Community area	i	–
Day	t	–
Trip count	*Yit*	Aggregated per community area per day per capita
Per capita income	γ_i	Aggregated per community area
Income group	Γ_i	Income split into high and low at median
Percent non-vehicle owners	ν_i	Non-vehicle owners per capita
Percent ethnic minority	μ_i	Ethnic minorities per capita
Walk score	ω_i	Walk score averaged per community area
Bike score	ρ_i	Bike score averaged per community area
Transit score	τ_i	Transit score averaged per community area
Stay-at-home order	σ_t	Before and during Stay-at-home order
Weekday	δ_t	Day of the week in 2020
Error	*Eit*	Regression error per day per community area

in the difference-in-difference formulations is the interaction terms between per capita income (continuous and grouped into high & low) and the stay-at-home order.

Regression Eq. 1: Continuous Income

$$log(Y_{it}) = \beta_1 log(\gamma_i) + \beta_2 log(\nu_i) + \beta_3 log(\mu_i) + \beta_4 log(\omega_i) + \beta_5 log(\rho_i) + \beta_6 log(\tau_i) + \beta_7 \delta_t + \beta_8 \sigma_t + \boldsymbol{\beta_9 \mathbf{log}(\gamma_i)\sigma_t} + E_{it} \tag{1}$$

Regression Eq. 2: Income groups

$$log(Y_{it}) = \beta_1 \Gamma_i + \beta_2 log(\nu_i) + \beta_3 log(\mu_i) + \beta_4 log(\omega_i) + \beta_5 log(\rho_i) + \beta_6 log(\tau_i) + \beta_7 \delta_t + \beta_8 \sigma_t + \boldsymbol{\beta_9 \Gamma_i \sigma_t} + E_{it} \tag{2}$$

The interaction between the continuous variable, the log of per capita income of the community area, and the stay-at-home order measures the incremental sensitivity to the stay-at-home order by income. Per capita income is a continuous variable, so the interaction term measures the rate of change in ridership per incremental income change. We cluster the errors at the community area level.

Equation 3 shows how to interpret the effect of a coefficient on a logged dependent variable. The interaction term between the log of per capita income and the stay-at-home order is continuous, therefore it represents the slope of difference. Eq. 3 yields the percent change the coefficient (β) has on the dependent variable when the dependent variable is logged.

$$\%\text{effect} = 100 \times (e^{\beta} - 1) \tag{3}$$

Data exclusion

Two community areas were excluded because they contain an airport creating an outlier. We are primarily assessing TNCs used by residents. Airport community areas have disproportionately high TNC ridership pre-pandemic given the per capita income of the community area because of the large volume of airport TNC riders that are not residents or workers in the community area that the airport is located in (See Appendix C for quantile-quantile plots with and without the two airport community areas included).

Raw TNC trip data includes each trip's origin and destination at the census tract and community area level. Some trips have all or partial location data excluded if the trip is the only trip in the geographical area in the 15-minute window in the dataset for rider location anonymity and privacy concerns. Census tracts are a smaller geographical designation than community areas, so trip origins are more often excluded from the data at the census tract level than at the community area level for privacy concerns. In February 2020 (before the stay-at-home order) 28% of census tract and 6% of community area origin points were excluded. In April 2020 (after the stay-at-home order) 88% percent of census tract and 9% of community area origin points were excluded for rider privacy. To mitigate the data exclusion issue we analyze a larger geographic region, aggregating to the community area, in which only 6% before the stay-at-home order and 9% after the stay-at-home order of data values are excluded for privacy. The community area level has both less overall data exclusion and a similar rate of data exclusion before and after the start of the stay-at-home order to minimize bias in data exclusion after the stay-at-home order (See Appendix D for a discussion of exogeneity). Most community areas contain census tracts of two or fewer income groups. Very few community areas contain census tracts across four or more income groups, signaling that income in the community area is still a granular enough designation to yield meaningful conclusions. See Appendix Efor a visual of data aggregation.

Five-fold cross validation

We use the 80% of the data not used in model formulation to conduct a five-fold cross-validation in order to verify that the model is not over-fit to the training data and accurately represents unseen data in the set. The five-fold cross-validation splits the data into five sections; it trains the model on four sections of data and tests the model on the remaining section. This is repeated until all five sections have been the test set. Both regression models are tested and the mean root mean squared error (rMSE) of the five-fold test is compared for the training set and the testing set for each regression. The first regression, Equation 1, contains an interaction term between the stay-at-home order and income as a continuous variable. The second regression, Equation 2 contains an interaction term between the stay-at-home order and income groups (split at the median into high and low).

Parallel trends

We assume the TNC trip trends by income are parallel leading up to the pandemic. To verify if the parallel trends assumption holds an event study using weeks away from the stay-at-home order (W_j) as a dummy variable in the interaction term is used in place of the stay-at-home order dummy for the second regression and the coefficient values are plotted relative to the coefficient value for the week of the stay-at-home order to ensure that parallel trends were followed leading up to the stay-at-home order. Equation 4 shows the regression formulation for the event study using a weeks-away from the stay-at-home order dummy (W_j) in the interaction term and no day-of-the-week dummy. The weeks-away dummy variable has a variable for each week leading up to and following the start of the stay-at-home order and assigns each trip as ‘1’ if the trip occurs in that week’s dummy variable and ‘0’ otherwise; No intercept is used because each trip falls into one week or another and an intercept would cause linear independence to be broken.

$$log(Y_{it}) = \beta_1 log(\gamma_i) + \beta_2 log(\nu_i) + \beta_3 log(\mu_i) + \beta_4 log(\omega_i) + \beta_5 log(\rho_i) + \beta_6 log(\tau_i) + \beta_7 W_j + \boldsymbol{\beta_8 \log(\gamma_i) W_j} + E_{it} \tag{4}$$

We also assume that the TNC trips would have continued at the same rate if the pandemic and stay-at-home order had not occurred. To verify the validity of this assumption, regression 2 is run on TNC trips for the previous year (2019) to verify that the trend noticed in 2020 is not due to seasonal changes, such as warmer weather or more daylight. The interaction term and stay-at-home order coefficient should be insignificant ($p > 0.05$) in the previous year if the trends are not due to seasonality, as we assume.

Origin–destination pair analysis

To better understand the TNC ridership trends that we see in the regression results, we conduct a posthoc analysis comparing the origin-destination pairs by income group (split at the median) for February 2020 (before the Stay-at-home order) and April 2020 (after the stay-at-home order) in absolute numbers and by hour of the day.

We assume that TNC riders conduct a closed loop each day (i.e. each person returns to their origin each day). Therefore, we would expect to see low-to-high-income area origin-destination (OD) pairs mirror high-to-low-income area origin-destination pairs. We aggregate OD pairs over a month, therefore we expect a similar departure and origin rate from each community area, assuming riders leave a given community area as frequently as they return to it. TNC inequities can be seen in OD pairs when asymmetrical flows occur between origins and destinations. We hypothesize that during the stay-at-home order, proportionally more TNC riders are essential riders, meaning, they are taking the ride because they require it for work or essential services like health care appointments and lack other modes of transportation to make their trip. We further hypothesize less choice ridership during the pandemic; choice riders have multiple options for transit and choose based on circumstances like weather, time of day, and convenience of each option. and the Transit Center have shown a reduction in public transportation ridership during the COVID-19 pandemic, leaving transit-dependent commuters who are disproportionately low-income, minority, and essential workers (Kantamneni 2020) (Transit Center 2020; Kantamneni 2020). We explore TNC trip OD-pairs by income group to understand if TNCs experienced similar shifts.

Figure 2 shows the data aggregation process for each dataset by time resolution and spatial resolution. TNC trip data shows each trip at the 15-minute interval with both census tracts and community areas of the origin and destination. The mobility information of the community (walk score, bike score, and transit score) at the neighborhood level are aggregated to the community area level and matched to each trip origin and destination. The demographic data from the American Community Survey (per capita income, percent non-vehicle ownership, and percent ethnic minority) are aggregated to the community area level using a population-weighted average from the census tract level. The trip origins and destinations are then split into high-income and low-income by the hour of the day that the trip started. This results in a trip count of origin-destination pairs for each possible income pairing by the hour (e.g., high-income area to high-income area, high-income area to low-income area, low-income area to high-income area, and low-income area to low-income area per hour).

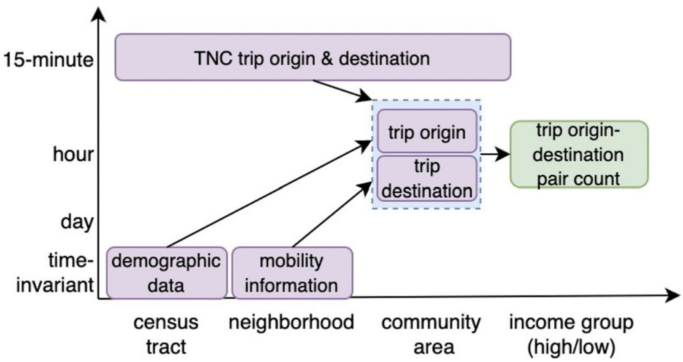

Fig. 2 Diagram of data aggregation process for posthoc analysis to combine each trip with the income group of its origin and destination by the hour

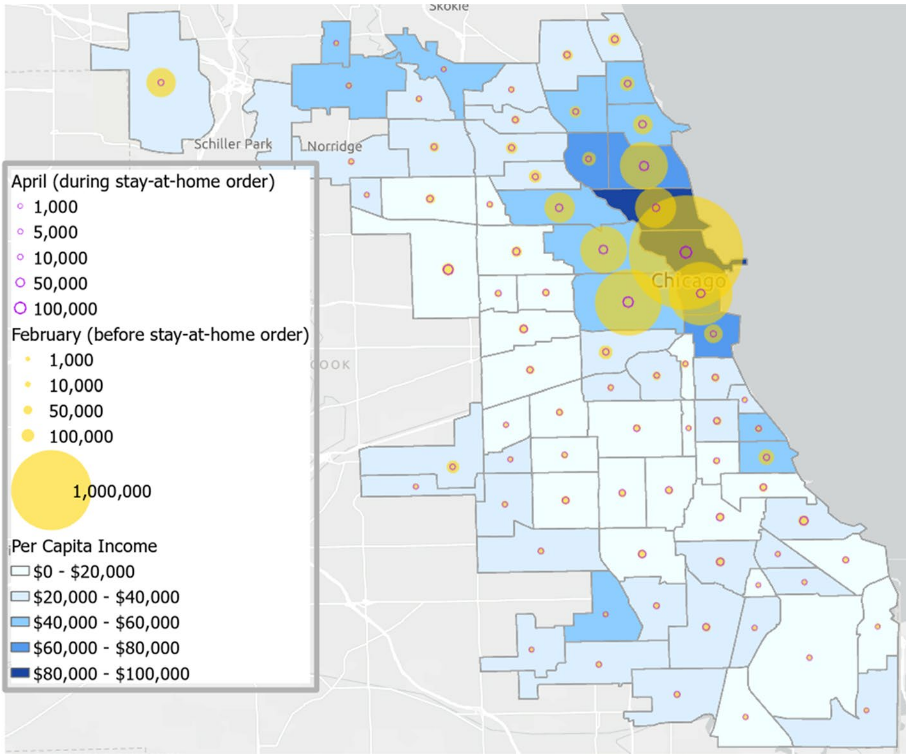

Fig. 3 Map of per capita income (grouped per $20,000) at the community area level with TNC trips for the month of February (before the stay-at-home order) and the month of April (after the stay-at-home order) as concentric circles for each community area

Results

We see a decrease in TNC trips across the entire Chicago area after the start of the stay-at-home order (84% decrease in total trips between February and April). However, some areas saw sharper decreases in trips compared to others. Figure 3 shows the change in trips before and during the stay-at-home order as well as the per capita income of each community area in blue. Trips for February (before the stay-at-home order) are proportional to size in yellow and trips for April are in purple. The downtown area, a high-income area had significant ridership prior to the stay-at-home order. Low-income areas largely had similar ridership before and during the pandemic, which can be seen by very little difference between the yellow and purple ridership circles. Across the map, during the stay-at-home order, levels of ridership were low but largely consistent across the community areas. One exception to pre-pandemic trends is the community areas that contain the airports (O'hare International and Midway) were relatively low-income areas with high pre-stay-at-home order ridership and therefore a significant drop. See Appendix F for a map of TNC trips before and after the stay-at-home order by percentage of ethnic minority residents.

Regression

Strict exogeneity is assumed for the interaction term of both regressions. We assume the error term is not correlated with the independent variables. We make this assumption because there is unlikely to be reverse causality between per capita ridership and the community area variables such as income, transit access, and demographics on the time-scale

Table 3 Coefficients for regression 1 (continuous income) & regression 2 (grouped income) in the interaction term. Robust standard errors clustered per community area

Independent variable	Regression 1: Continuous income Estimate (Robust std. error)	Regression 2: Income groups Estimate (Robust std. error)
Sunday	– 18.37*** (1.56)	– 13.41*** (1.53)
Monday	– 18.34*** (1.56)	– 13.39*** (1.53)
Tuesday	– 18.33*** (1.56)	– 13.38*** (1.54)
Wednesday	– 18.28*** (1.56)	– 13.32*** (1.53)
Thursday	– 18.37*** (1.56)	– 13.25*** (1.53)
Friday	– 18.06*** (1.56)	– 13.10*** (1.53)
Saturday	– 18.13*** (1.56)	– 13.17*** (1.53)
Log (per capita income)	1.04*** (0.13)	0.57*** (0.14)
Log (percent no vehicle)	0.50*** (0.048)	0.50*** (0.048)
Log (transit score)	1.89*** (0.27)	1.85*** (0.30)
Log (walk score)	– 0.40 (0.27)	– 0.44 (0.28)
Log (bike score)	– 0.06 (0.36)	– 0.0022 (0.39)
Log (percent ethnic minority)	0.33** (0.11)	0.33** (0.11)
Stay-at-home order	9.40*** (0.66)	– 1.63*** (0.088)
Income group (low)	–	– 0.35** (0.12)
Income group (low):stay-at-home order	–	0.83*** (0.10)
Log (per capita income):stay-at-home order	– 1.04*** (0.064)	–

Signif. codes: ***$P(0)$, **$P(\leq 0.001)$, *$P(\leq 0.01)$, $P(\leq 0.05)$

of the analysis; meaning that the TNC ridership in an area during COVID-19 did not impact the income or demographics of the area.

Table 3 shows the coefficients, clustered robust standard errors, t-values, and probability of being greater than the t-value for the first regression (with income as a continuous variable in the interaction term). All variables except the walk score and bike score are statistically significant ($p<0.05$). The interaction term between the log of per capita income and the stay-at-home order is continuous, therefore it represents the slope of difference. Between two neighborhoods, if one has 10% higher per capita income, the stay-at-home order is correlated with a 10.4% additional decrease in trips. Income as a continuous variable within the interaction term gives an indication of the degree of impact that income has on the rate of change of ridership during the stay-at-home order.

The dependent variable for the second regression in Table 3 (Trips per capita per community area per day) is logged; therefore, the interaction term between income groups and the stay-at-home order shows that high-income community areas have a 56% greater decrease in the number of TNC trips origins compared to low-income community areas. High-income areas had more trips per capita prior to the stay-at-home order; however, high-income areas had fewer trips after the pandemic with high-income areas decreasing trips by 80% (Fig. 4). Riders taking trips from low-income areas continued depending on TNCs at higher rates than high-income community areas to fulfill necessary trips. Trips from low-income areas showed significantly less ability to avoid TNC trips during the stay-at-home order. The high-income community areas had higher per capita ridership before the stay-at-home order and had a bigger decrease in ridership during the stay-at-home order than low-income areas.

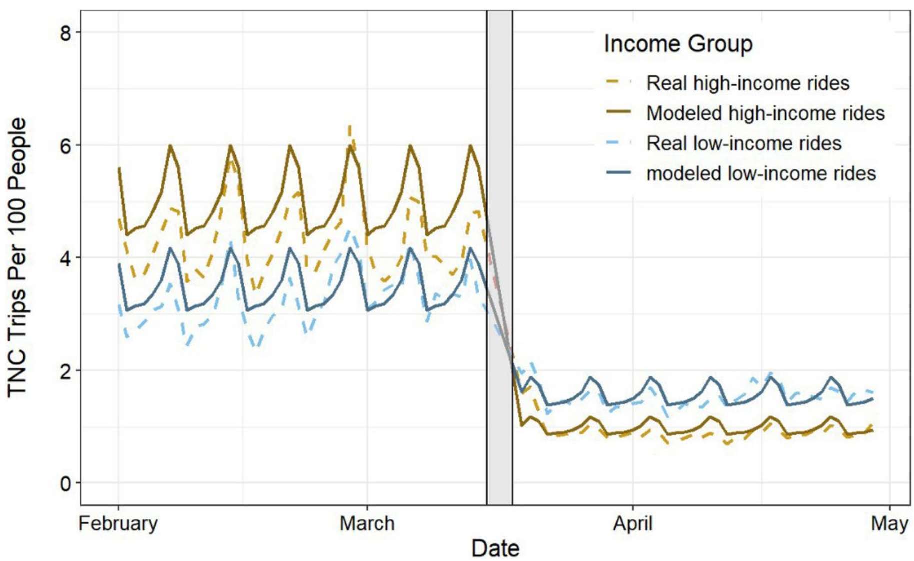

Fig. 4 Modeled trips per capita for regression 2 (income grouped into high and low) and actual trips from Feb. 1st to April 29th 2020 with a grey zone indicating the transitional period between the CDC guidance and the beginning of the stay-at-home order beginning March 18th

Table 4 Root mean squared error (rMSE) for training and testing data for the five-fold cross-validations

Model	Training rMSE	Testing rMSE
Regression 1: continuous income	0.0045	0.0089
Regression 2: income groups	0.0050	0.0099

Model 1: interaction term includes income as a continuous variable and the stay-at-home order

Model 2: interaction stay-at-home order term between income (high & low)

Five-fold cross validation

We use 80% of the data to conduct five-fold cross validation to ensure the model is not over-fitted to the data. Both regressions had a testing root mean squared error (rMSE) that was just under double the training rMSE. Table 4 shows that overall, the rMSEs for both the training and testing set were low (less than 1% in all cases). Regression 1 (income as a continuous variable in the interaction term) performs slightly better for both the training and testing data, indicating it might be a better fit than regression 2 (income as a discrete variable). The higher testing rMSE than training rMSE for both regressions indicates that the models may be slightly over-fitted to the data. However, the rMSE values are low for both sets for testing, implying a good overall fit, despite some over-fitting.

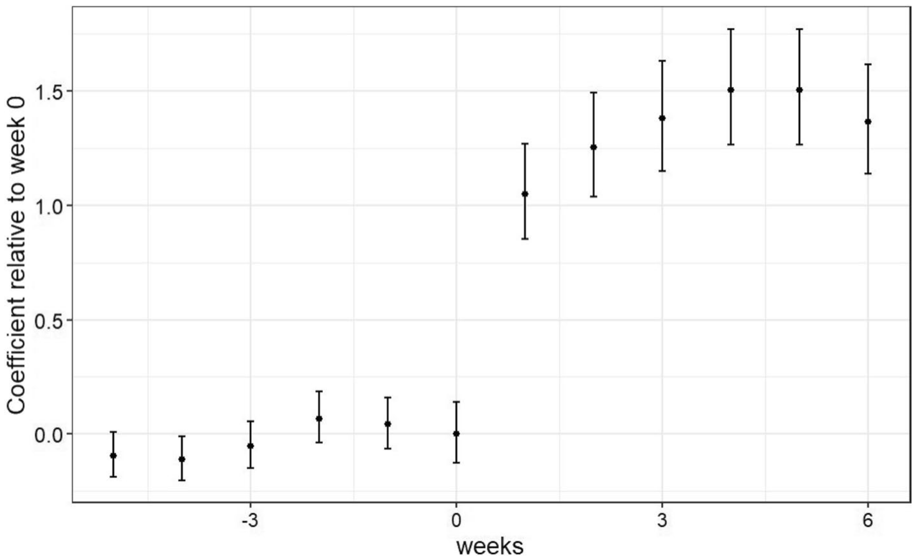

Fig. 5 Event study of parallel trends with the coefficient for the interaction term for each week normalized to the week the stay-at-home order is put in place (week 0) with the 95% confidence interval on the coefficients

Parallel trends

We conduct an event study using each of the six weeks before and after the stay-at-home order as dummy variables to check the parallel trends assumption. In Fig. 5, the coefficients represent the interaction term between each week and the income group normalized to the week the stay-at-home order is mandated (week 0). The bars represent the 95% confidence interval on each coefficient. If the parallel trend assumption holds, the weeks leading up to the stay-at-home order should have no difference in their coefficient compared to the week of the stay-at-home order, leading to a value of zero. Indeed, in the event study, the weeks leading up to the stay-at-home order hover around zero (−0.012 to 0.17) and after the stay-at-home order the coefficient values jump significantly (1.15 to 1.60) with confidence intervals well above zero (0.95 at the lowest bound), indicating the parallel trends assumption holds.

Using the full set of data, we check the parallel trends assumption for 2020 compared to the previous year. We assume the parallel trends assumption holds, that is if the Covid-19 pandemic and stay-at-home order had not occurred, ridership by income would remain parallel. Our regressions do not account for seasonal changes in ridership. Given the contrast before and after the stay-at-home order (week 0), we assume the changes observed in the event study are explained by the stay-at-home order and not by changes in the season. Figure 6 shows the trips per day broken up by $20,000 based on the trip origin from February 1st to April 29th (a) 2019 and (b) 2020. We group by $20,000 to show differences in ridership by income within the region without plotting a line for each community area. Early 2019 did not have a stay-at-home order and is pre-pandemic, so any seasonal effects would be visible within the 2019 trips per day. 2019 lacks visible seasonal trends for February to April of 2020, indicating that the observed decline in ridership in 2020 is not seasonal. See Appendix G for the regression table for 2019 showing the insignificance ($p > 0.05$) of the stay-at-home order interaction

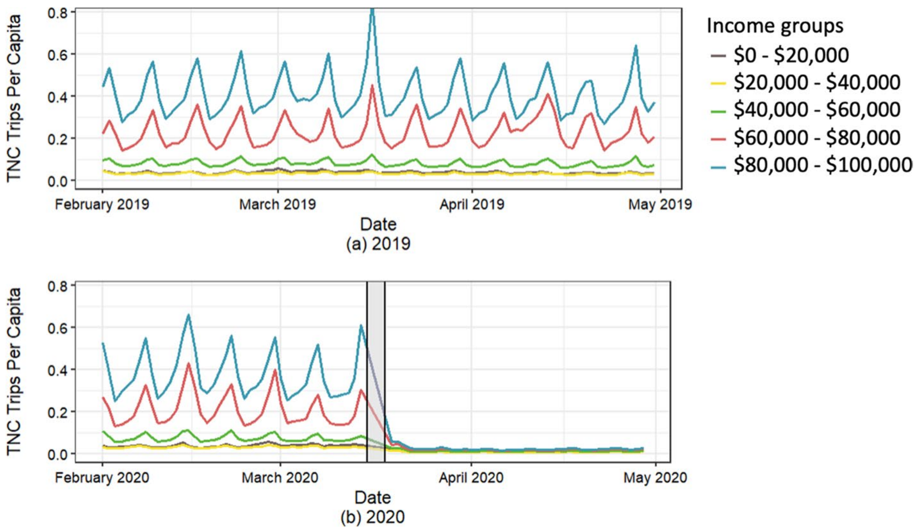

Fig. 6 Ridership each day by income bracket in (a) 2019 and (b) 2020 demonstrating the parallel trends throughout the season for 2019 with rides grouped by $20,000 (<$20,000, $20,000—$40,000, $40,000—$60,000, $60,000—$80,000, $80,000—$100,000)

term for the previous year. We re-run regression 2 for the previous year's trips, setting the stay-at-home order to March 18th, 2019 (the previous year) to verify there is insignificance to that date, yearly. We find no significance to the post-March 18th period for 2019, indicating that the parallel trends assumptions hold: if the pandemic and stay-at-home orders had not occurred, ridership would have continued at a similar rate.

Origin–destination Pairs

We look at the income group (low or high) of each trip origin and destination community area to compare the net trips by per capita income. Table 5 shows the aggregate difference between high-to-low-income area trips, and low-to-high-income area trips for February 2020 (before the COVID-19 pandemic) and April 2020 (during the stay-at-home order). In February, we see 32% more trips from high-to-low-income areas (353,000 trips) than trips from low-to-high-income areas (239,000 trips). This difference indicates that a portion of TNC riders are multi-modal (only take TNCs in one direction). This could indicate a limited supply of TNCs in low-income areas or a preference for TNCs in the evening for low-income residents returning from high-income areas. From February to April 2020 we see an 84% drop in trips following the stay-at-home order. Meanwhile, we see only 8% more high-to-low-income trips than low-to-high-income trips. In April, the 16% of baseline trips that still occur show greater mirroring of origin-destination pairs from low-to-high and high-to-low (8% difference). The symmetry between low-to-high and high-to-low indicates during the stay-at-home order, TNC users take a TNC in both directions and display less choice-ridership compared to pre-pandemic riders (i.e. potentially lack other options for their trips).

The peak of trips occurs earlier in the day in April, with total TNC ridership peaking at 3 PM compared to 6 PM in February prior to the stay-at-home order. April 2020 shows overall less of a disparity between trips from high-to-low-income areas and trips from low-to-high-income areas (8% difference). This may indicate less choice-ridership in low-income areas during the stay-at-home order; fewer trips are only completed by a TNC in one direction, potentially showing that fewer TNC trips had alternative options for those trips.

Figure 7 shows the trips from high-to-low-income community areas and low-to-high-income community areas aggregated by the hour for February 2020 and April 2020. In February, high-to-low-income trips surpassed low-to-high-income trips for 23 out of 24 hours of the day. We find the largest gap at 6 pm with 43% more high-to-low income trips. This asymmetry indicates that TNC riders are only taking TNCs in one direction. This most likely reflects differences in supply constraints either of public transit or TNC drivers. Figure 7b of April (during the stay-at-home order) shows a greater degree of mirroring: the low-to-high income trips appear to mirror the high-to-low income trips (53% more low-to-high income area trips at 6 AM followed by 41% more high-to-low income area trips at 5 PM).

Conclusion

Our paper demonstrates a method for evaluating the impact of policy on TNC ridership across different demographic groups and the impact policies have on overall TNC usage. We use the COVID-19 pandemic as a natural experiment to assess the ability to

Table 5 Disparity of trips that originate in a high-income community area to end in a low-income compared to trips that originate in a low-income community area and end in a high-income community area

Month	High-to-high income	Low-to-low income	High-to-low income	Low-to-high income	Difference between high-to-low and low-to-high (%)
February	2,256,943	2,614,621	1,764,224	1,193,703	32
April	484,085	476,096	158,961	146,314	8

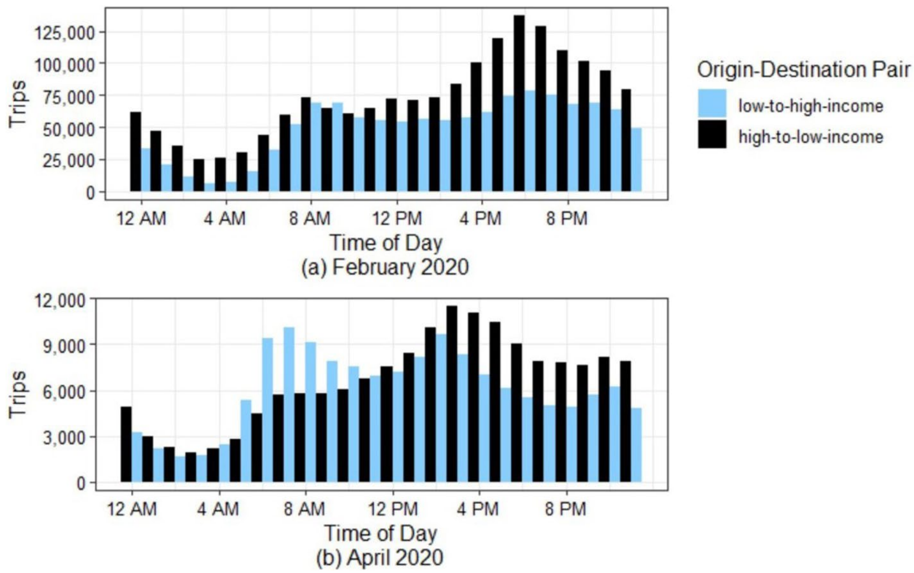

Fig. 7 TNC trips in Chicago aggregated by the hour of the day for trips going from high-to-low-income community areas and for trips going from low-to-high-income community areas for **a** February 2020 (pre-stay-at-home order) and **b** April 2020 (during stay-at-home order). **a** February and **b** April have different y-axis scales because (pre-pandemic) February had an order-of-magnitude greater number of trips per hour

avoid trips in response to policies and mandates. Low-income community areas have less ability to work from home during a stay-at-home order and fewer transportation choices like a private car or the ability to walk/bike to work (TransitCenter 2020; Kantamneni 2020). We focus here on TNC trips (e.g., Uber, Lyft, and Via) to see how different groups respond to policies that impact ridership. This paper measures the difference in sensitivity of TNC ridership between low and high-income community areas, using Chicago's first stay-at-home order in March 2020 as a natural experiment. We set up two difference-in-difference style regressions looking at the interaction between income and the stay-at-home order. All TNC riders received the treatment of the stay-at-home order; however, different areas had heterogeneous responses in their behavior in accordance with the policy.

Our paper reveals the essentialness of TNC trips by income to give an indication of which areas and groups may have the most difficult time avoiding trips during transportation shocks (e.g., health crises, infrastructure failures, political unrest, or severe weather). This can provide an indication of where resources should be allocated during rare events when some or all transportation is disrupted. TNCs are often considered a luxury good due to their primary riders in pre-pandemic conditions residing in high-income areas with multiple modal options (Grahn et al. 2020; Barajas and Brown 2021). We uncover that riders from low-income areas have a higher degree of unavoidable trips during the stay-at-home order. TNCs may be a service of convenience for most pre-pandemic riders; however, riders from low-income areas may be disproportionately dependent on TNCs during periods of disruption and disproportionately dependent on public options for transportation. When making policy decisions on TNCs for tax structures and availability, it is valuable to identify TNC-dependent riders among all TNC riders. Some TNC providers include transportation access in their mission statement (Uber 2022), although providers use surge pricing during times of peak demand to maximize profit. This surge pricing can overlap with essential trips, which further increases inequities during times of high essential ridership. Understanding which groups have more inelastic demand for TNCs as an essential service could help cities design programs such as providing vouchers or supplementing the cost of rides at hours when operating public transit at high frequency is not economical.

We assume that during the first six weeks of the stay-at-home order, most TNC trips were necessary (e.g., groceries, work, medical) and could not be completed with a private vehicle. If any TNC trips were non-essential during the stay-at-home order (i.e., to fulfill a non-essential trip), we cannot differentiate these trips from essential trips with the data available. Additionally, we assume that the stay-at-home ridership, which is an order of magnitude lower in the trip count, contains a subset of pre-pandemic riders, rather than a new set of riders. Not every trip that departs from a low-income area contains a low-income passenger. However, the city of Chicago bases TNC tax rates on the trip origin and destination, giving precedence to location-based analysis over passenger-based analysis (City of Chicago 2020c). Thus, the origin and destination are still useful as proxy metrics for making policy decisions. Additionally, the TNC data set does not indicate if the same passenger made multiple trips (or round trips). We investigate the symmetry of inbound and outbound trips to gain a sense of the degree of choice-ridership and to identify asymmetries in TNC origin-destination pairs. Passenger-level information is not publicly accessible due to privacy concerns. While the socioeconomic and demographic characteristics of each rider can give a greater sense of equity impacts, we are unable to determine rider-specific demographic impacts on usage rates. We used the characteristics of a trip's origin or destination as a proxy for rider demographics. This allows analysis of millions of rides across the entire city for a depth and completeness of data that rider-specific surveys cannot capture.

Our findings may indicate that high-income community areas are potentially more able to adapt to policies affecting TNCs, such as stay-at-home orders because of the nature of their jobs or factors like higher private vehicle ownership or better public transit infrastructure. This higher sensitivity may indicate differences in the essentialness of TNC ridership. Although high-income community areas have higher average ridership prior to the stay-at-home order, low-income community areas have higher ridership during the stay-at-home order and less sensitivity to the policy overall. Additionally, COVID-19 is no longer at its peak (City of Chicago 2020b), there is persistent work-from-home among office workers (Pawar et al. 2021) indicating that trends in ridership may carry over to post-pandemic transportation.

TNC ridership reflects supply in equilibrium with demand (i.e., there was a driver willing and able to provide a trip and a passenger desiring that trip in close enough proximity). We acknowledge the limitation that a decrease in ridership may also reflect a decrease in the supply of TNC drivers available to provide rides. According to a survey by Rideshare Guy of 1,000 TNC drivers nationwide, 58% of drivers surveyed stopped driving during initial stay-at-home orders, which is a smaller decrease than overall trips in Chicago (84% decrease), indicating that there is still some persistence of driver availability to meet demand (Goldstein 2020). One limitation of the TNC data made available by the city of Chicago is the lack of data on latent demand (i.e., any trips that would have occurred if there were a driver available), which makes it difficult to understand how driver availability affected the decrease in trips that occurred during the pandemic. We encourage cities that publish data on ridership to also publish data on requested trips that were not fulfilled. Data on wait times and latent demand would provide an opportunity to further research how equitable the distribution of supply is for demand and if there is bias in wait time or rate of unfulfilled trips by location.

Riders from low-income areas demonstrated a greater dependence on TNCs during the stay-at-home order than riders coming from high-income areas. This could be due to the higher persistence of low-income area TNC ridership when only essential trips were permitted, or it could indicate that some riders from low-income areas chose TNCs over public transportation during the stay-at-home order due to decreased service or perceptions of safety (Hanig et al. 2023). The relatively high rate of essential trips from low-income areas reveals an inability to complete their ride via a private mode of transportation (e.g., walking, biking, or a private car) or reflects the inability to avoid their trip entirely (e.g., work from home, receive grocery deliveries). Finding where TNC ridership remains relatively high while travel for non-essential trips is limited provides an opportunity for policymakers to better understand which areas are most dependent on public options for travel and least able to avoid their trips. This can provide insight into where to dedicate public transportation resources (e.g., additional bus routes) post-pandemic or during future transportation shocks. When TNCs serve as a substitute during transportation shocks while transportation service is limited, it is valuable for policymakers to understand which areas are most dependent on public options for travel and least able to avoid their trips. When TNCs serve as a complement to public transportation, persistent TNC ridership during the stay-at-home order can reveal where TNCs are providing an essential service, indicating potential latent demand for public transportation. One limitation of the TNC data from the city of Chicago is the lack of public transportation ridership data (with origin-destination pairs) at a granular enough time scale to be coupled with TNC ridership. If TNC and public transportation origin-destination pairs are both made available, the degree to which riders from low-income areas persisted in riding TNCs during the stay-at-home order vs. began riding TNCs when public transportation options were limited could be disaggregated. This highlights a potential area for further research by allowing TNC origin-destination pairs with higher-than-expected ridership given demographics to be compared against the public transportation options to identify new potential routes.

Much of the current literature states that the majority of TNC riders are high-income and many have alternatives to TNCs, making TNCs a luxury and not a necessity (Grahn et al. 2020; Barajas and Brown 2021; Jin et al. 2019). However, we find that pre-stay-at-home order trips were disproportionately from high-income areas, and when mandated to stay home a majority of trips departed from low-income community areas. This indicates that although in normal conditions trips from low-income areas are a minority of overall TNC trips, these trips are more likely to be essential than the trips coming from high-income areas. Understanding which areas have persistent ridership during the stay-at-home order could inform where new public transportation routes could be needed, where existing routes need more frequent service, or where resources need to be allocated during periods of transportation disruption. Areas with high rates of essential ridership could also inform tax structures for TNCs; the city of Chicago has previously structured TNC taxes based on the origin and destination of TNC trips during key times of the week to reduce congestion (by increasing the TNC tax in high-congestion areas) (City of Chicago 2020c). Tax rates for TNCs could be optimized by not taxing TNCs that are likely to be serving an essential trip while increasing the cost in locations and times where the trip is most likely to be avoidable and a source of congestion. Finding which areas have a higher dependence on TNCs can inform policy, infrastructure, and transit decisions to improve access and mobility equity for the most transit-dependent riders.

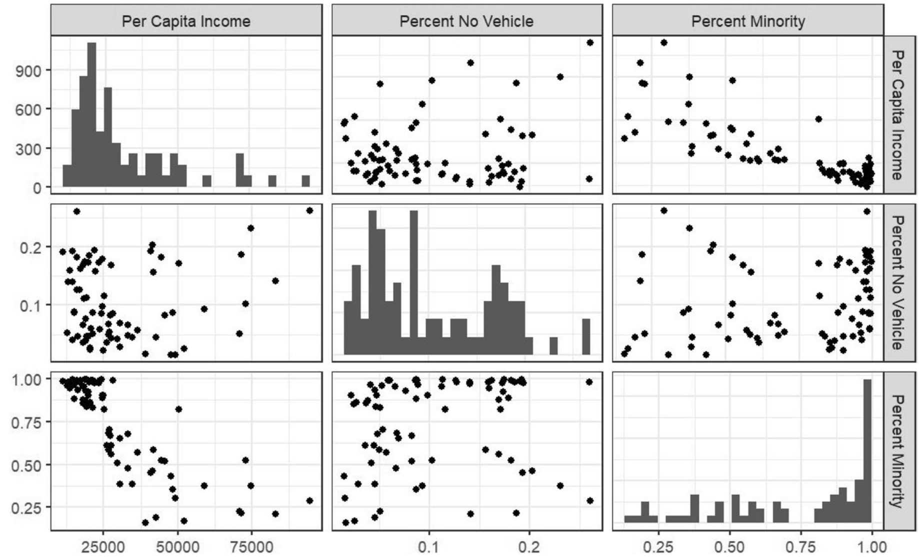

Fig. 8 Pair plot of community area demographics for trips using 20% of data for model formulation

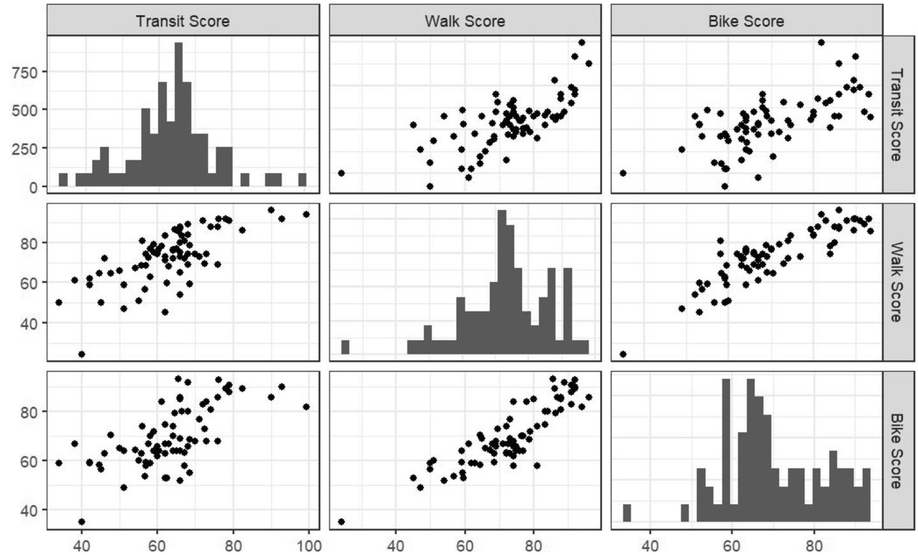

Fig. 9 Pair plot of community infrastructure characteristics for trips using 20% of data for model formulation

Appendix A: Distribution and descriptive statistics

The distribution of variables compared to each other are graphed in Figs. 8 and 9. If two variables are highly correlated, it may break the co-linearity Gauss-Markov assumption because the variables would not be independent. Overall there is not a very distinct trend in the correlation between the demographic correlations, except per capita income and percent minority show some negative correlation with heteroskedasticity. Walk score, bike score, and transit score are highly correlated, which may be why they had lower overall significance in both regressions.

Appendix B: Model formulation

We use a Box-Cox transformation of the dependent variable (Fig. 10) to formulate the dependent variable relationship. The Box-Cox transformation gives an exponentiation fit of $\lambda = -0.02$ (close to zero) indicating that a log transformation of the dependent variable (Trips per capita per community area per day, Y_{it}) is a good fit. See Appendix E for correlation plots and histograms of the independent variables. Equations 5–9 show pseudo-code of how the box-cox function from the R library 'Mass' is used to find the lambda value and therefore the best-fit transformation for the dependent variable.

Pseudo-code of the Box-Cox transformation

$$\begin{aligned}
\text{regression 1 output} = \text{linear regression (log(TNC trips per day per community area)} \\
\sim \text{log(income)} + \text{log(non} - \text{vehicleownership)} + \text{log(percent ethnic minority)} \\
+ \text{log(walk score)} + \text{log(bikescore)} + \text{log(transit score)} + \text{day of the week dummy variable} \\
+ \text{stay} - \text{at} - \text{home order dummy variable} + \text{log(income)} \times \text{stay} - \text{at} - \text{home order dummy variable} \\
- 1, \text{data} = \text{TNC trip data sampled for the first 20\% and including trips in community areas)}
\end{aligned}$$

$$\begin{aligned}
\text{box} - \text{cox output for regression 1} &= \text{Box} - \text{Cox transformation(regression 1 output)} \\
\text{lambda for regression 1} &= \text{maximum(box} - \text{cox output for regression 1)}
\end{aligned}$$

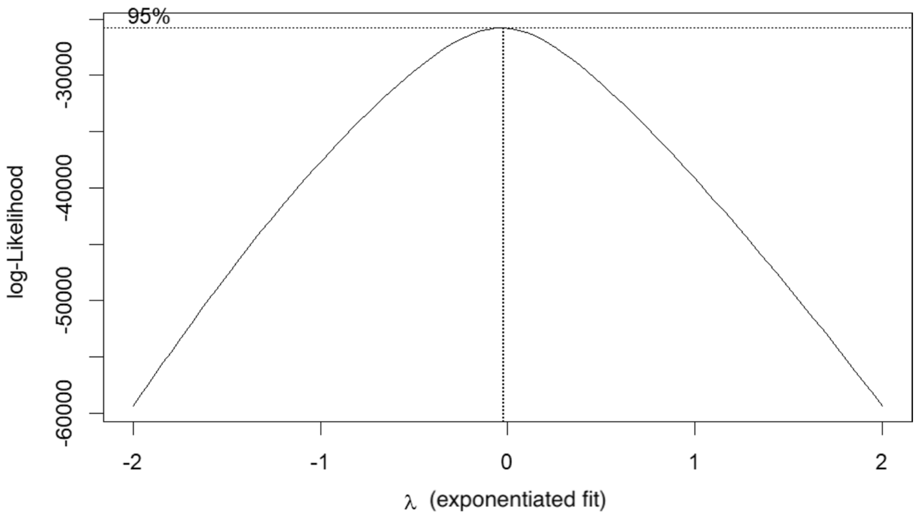

Fig. 10 Box-Cox transformation of the dependent variable, Trips per community area per day, indicating that a log transformation is an appropriate transformation ($\lambda = -0.02$)

$$
\begin{aligned}
\text{regression 2 output} = {} & \text{linear regression (log(TNC trips per day per community area)} \\
& \sim \text{log(income)} + \text{log(non} - \text{vehicle ownership)} + \text{log(percent ethnic minority)} \\
& + \text{log(walkscore)} + \text{log(bikescore)} + \text{log(transitscore)} + \text{day of the week dummy variable} \\
& + \text{stay} - \text{at} - \text{home order dummy variable} + \text{income} \times \text{stay} - \text{at} - \text{home order dummy variable} \\
& - 1, \text{data} = \text{TNC trip data sampled for the first 20\% and including trips in community areas)}
\end{aligned}
$$

$$
\begin{aligned}
\text{box} - \text{cox output for regression 2} &= \text{Box} - \text{Cox transformation(regression 2 output)} \\
\text{lambda for regression 2} &= \text{maximum(box} - \text{cox output for regression 2)}
\end{aligned}
$$

We run the gam function in R from the ‘gam’ package to test the transformation with a spline on each fixed effect to model each non-binary independent variable (Fig. 11). The spline transformation plots indicate that a log transformation is an appropriate fit for per capita income, percent non-vehicle ownership, percent ethnic minority, walk score, bike score, and transit score.

Appendix C: Airport exclusion

The city of Chicago has 77 community areas and two airports (O’Hare and Midway). The airports have disproportionately high TNC trips for their community-level metrics such as per capita income and the walk-score of the neighborhood because the majority of TNC trips in these two community areas are from the airport by non-residents of the community area. Figure 12 shows the jackknife residuals and quantile-quantile plots for the first regression for the first 20% of data used for model formulation. The qqplot with airports included shows significant deviation due to the airports having disproportionate

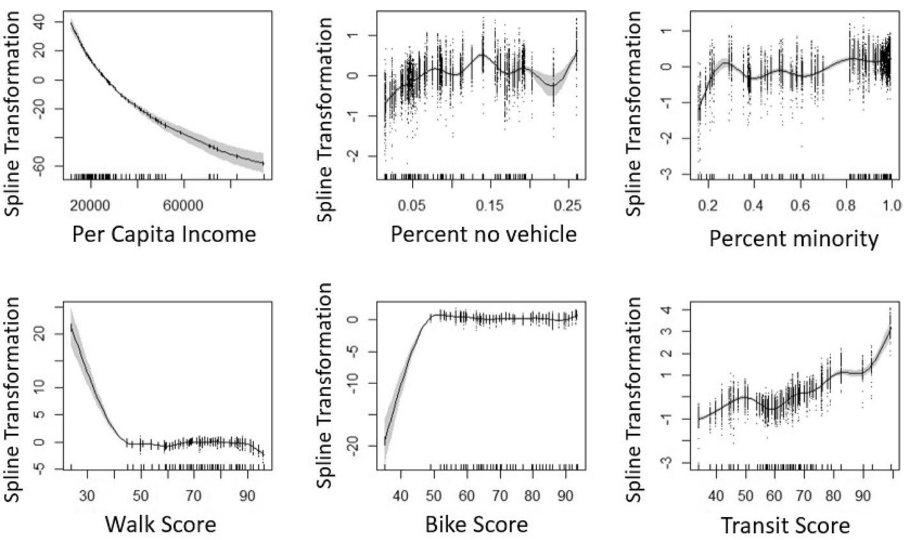

Fig. 11 Spline transformation plot for each independent variable indicating which (if any) transformations for the independent variables would be appropriate

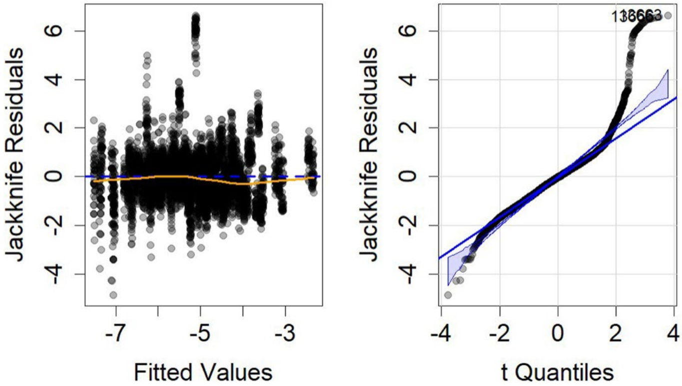

Fig. 12 Quantile–quantile plot showing the outliers from the initial 20% of data showing that the two community areas with airports had a disproportionately high number of trips prior to

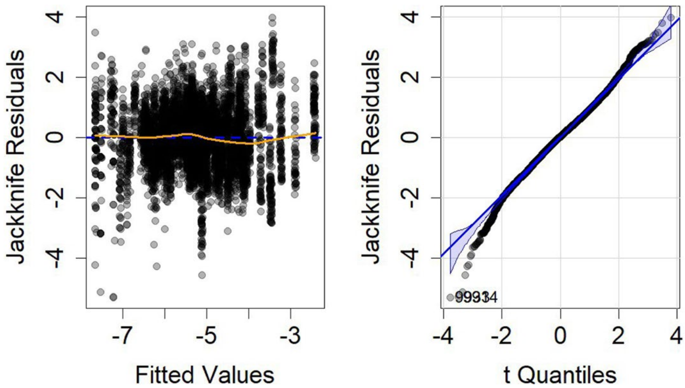

Fig. 13 Quantile–quantile plot with airport-containing community areas removed

trips. Figure 13 shows the quantile-quantile plot for the 20% of data used in the model formulation while excluding the two community areas that contain airports. The quantile-quantile plots no longer have the significant upward deviation present in Fig. 12. The outliers fitted into two categories; they either occurred in community areas with an airport (the trip count was higher than expected for the given community area demographics) or the trip count was exceptionally low. The latter is probably due to the biased way that data gets thrown out of the dataset in the case that there is only one trip in a community area in a 15-minute window, the location is not included. This makes community areas and days with low ridership appear to have even lower ridership because trips are not recorded as a part of that community area's total trip count that day.

We include pseudo-code for modeling quantile-quantile plots for the first 20% of the data (for model fitting) for airport exclusion. We show the code for regression 1 when run including community areas containing airports included to create Figure 12a and b and we show the code for regression 1 excluding community areas containing airports for creating Fig. 13.

Pseudo-code of plotting Jackknife residuals and quintile-quintile plots

$$
\begin{aligned}
\text{regression output with airports} = \text{linear regression (log(TNC trips per day per community area)} \\
\sim \log(\text{income}) + \log(\text{non} - \text{vehicle ownership}) + \log(\text{percent ethnic minority}) \\
+ \log(\text{walks core}) + \log(\text{bikescore}) + \log(\text{transit score}) + \text{day of the week dummy variable} \\
+ \text{ stay} - \text{at} - \text{home order dummy variable} + \log(\text{income}) \times \text{stay} - \text{at} - \text{home order dummy variable} \\
- 1, \text{data} = \text{TNCtrip data sampled for the first 20\%and including trips in community areas with airports)}
\end{aligned}
$$

$$
\begin{aligned}
&\text{figure 12a} = \text{plot(x} = \text{vector(fitted values of linear regression(regression output with airports)),} \\
&\quad \text{y} = \text{vector(studentized residuals(regression output with airports))} \\
&\quad\quad + \text{plot(y} = 0, \text{line color} = \text{blue)} + \text{plot(x} = \text{locally weighted scatter plot smoothing} \\
&\quad\quad \text{(vector(fitted values of linear regression (regression output with airports)),} \\
&\quad \text{y} = \text{vector(studentized residuals(regression output with airports)),} \\
&\quad\quad \text{linecolor} = \text{orange)}
\end{aligned}
$$

$$
\begin{aligned}
&\text{figure12b} = \text{quantile} - \text{quantile plot function(vector form(regression output with airports))} \\
&\quad \text{regression output without airports} = \text{linear regression (log(TNC trips per day per community area)} \\
&\quad \sim \log(\text{income}) + \log(\text{non} - \text{vehicle ownership}) + \log(\text{percent ethnic minority}) \\
&\quad + \log(\text{walkscore}) + \log(\text{bikescore}) + \log(\text{transitscore}) + \text{day of the week dummy variable} \\
&\quad + \text{stay} - \text{at} - \text{home order dummy variable} + \log(\text{income}) \times \text{stay} - \text{at} - \text{home order dummy variable} \\
&\quad - 1, \text{data} = \text{TNC trip data sampled for the first 20\% and including trips in community areas} \\
&\quad \text{without airports)}
\end{aligned}
$$

$$
\begin{aligned}
&\text{figure13a} = \text{plot(x} = \text{vector(fitted values of linear regression(regression output without airports)),} \\
&\quad \text{y} = \text{vector(studentized residuals (regression output without airports))+} \\
&\quad\quad \text{plot(y} = 0, \text{linecolor} = \text{blue)} \\
&\quad\quad + \text{plot(x} = \text{locally weighted scatter plot smoothing(} \\
&\quad\quad \text{vector(fitted values of linear regression(regression output without airports)),} \\
&\quad \text{y} = \text{vector(studentized residuals(regression output without airports)),} \\
&\quad\quad \text{linecolor} = \text{orange)}
\end{aligned}
$$

$$
\text{figure 13b} = \text{quantile} - \text{quantile plot function(vector form(regression output without airports))}
$$

Appendix D: Exogeneity

Contemporaneous exogeneity

The residuals were slightly homoscedastic. This indicates that the zero conditional mean of the error assumption may be violated. Some of this can be understood because of the biased way within low-trip count community area-day pairs had more trips thrown out, biasing the trip count to be lower than expected.

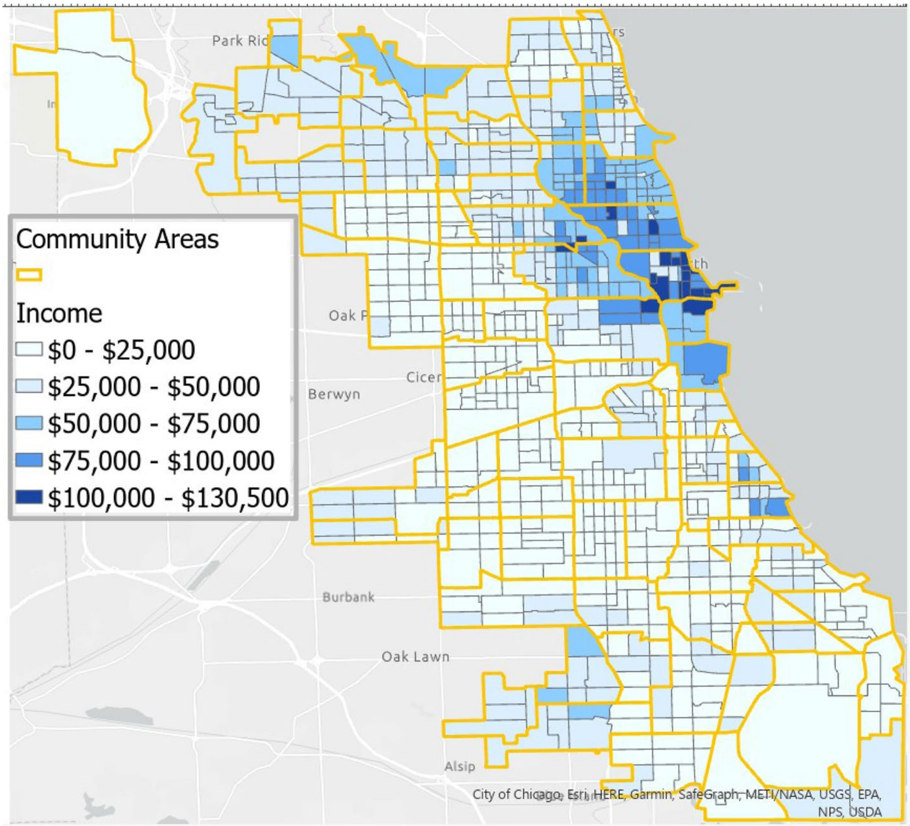

Fig. 14 Map of per capita income at the census tract level in Chicago with an overlap of community area boundaries in orange. Census tracts are grouped by every \$25,000 (e.g., < \$25,000, \$25,000-\$50,000, \$50,000-\$75,000, …, ≥ \$100,000)

Forward and backward exogeneity

For difference-in-difference style regression, a very important question related to causality is if the policy caused the change that is observed. Figure 4 shows an almost step-function appearance right at the stay-at-home order; how ever, it is not clear that the change is due exclusively to the order. If the 'treatment' is thought of as the pandemic itself with the stay-at-home order representing the start of concern for the pandemic locally, then the parallel trends assumption holds. To understand if the response was due to COVID-19 presence in the city or due to policies and administrative guidance, we compare case counts to policy dates. The COVID-19 Cases in Chicago in remained very low throughout the three months observed; however, TNC ridership fell steeply between March 15th and March 18th when the stay-at-home order went into place and the CDC recommended avoiding large groups, indicating that ridership was more responsive to policy than to actual case risk for the early months of the pandemic.

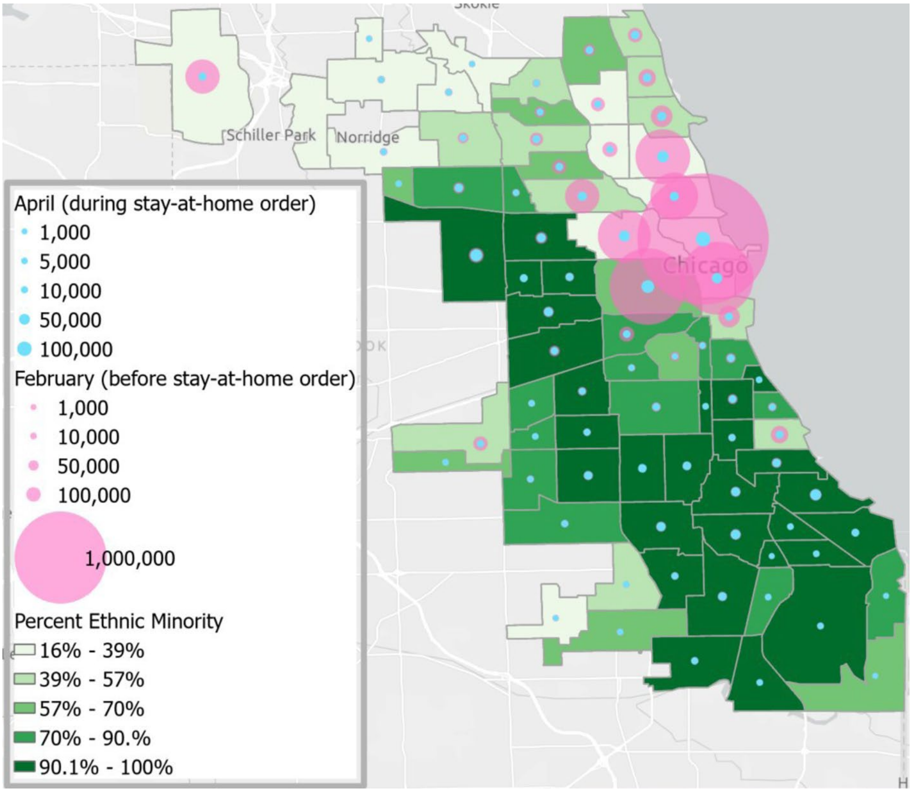

Fig. 15 Map of percent of each community area that is an ethnic minority with TNC trips for February (before the stay-at-home order) and April (during the stay-at-home order) as concentric circles for each community area

Appendix E: Data aggregation

We group data at the community area level into 77 community areas in the city of Chicago from 866 census tracts due to avoid data exclusion from TNC trips. We find a weighted average of per capita income at the community area level from the census tract level as outlined in the Methods section. Figure14 shows the per capita income at the census tract level with the boundaries of community areas highlighted in orange. The majority of community areas are highly homogeneous at the census tract level (i.e., two or fewer income groups of census tracts per community area), indicating that grouping by census tract does not lose a high degree of economic information.

Appendix F: Map of percent ethnic minority ridership

Percent ethnic minority and per capita income are highly correlated, which is why the regressions have an interaction term for the stay-at-home order and income, but do not add an interaction term between stay-at-home order and percent ethnic minority (see Appendix

Table 6 Table of coefficients for regression 2 for 2019 TNC trips with errors clustered at the community area

Independent Variable	2019 Estimate	Robust standard error
Sunday	−20.45***	−2.2
Monday	−20.47***	−2.21
Tuesday	−20.45***	−2.21
Wednesday	−20.43***	−2.21
Thursday	−20.35***	−2.21
Friday	−20.19***	−2.21
Saturday	−20.19***	−2.2
log(per capita income)	0.81***	−0.17
log(percent no vehicle)	0.42***	−0.07
log(transit score)	2.15***	−0.35
log(walk score)	−0.39	−0.43
log(bike score)	0.69	−0.47
log(percent ethnic minority)	0.14	−0.14
Stay-At-Home Order	0.07***	−0.008
Income Group (low)	0.04	−0.13
Income Group (low):Stay-At-Home Order	0.01	−0.01

Signif. codes: ****P*(0), ***P*($\leq$0.001), **P*($\leq$0.01),. *P*($\leq$0.05)

A and Fig. 7 for a pair plot of variable correlations). Figure 15 shows the map of the percent ethnic minority population by community area with the number of trips for February and April. Community areas with fewer ethnic minorities generally saw a sharper decrease in ridership, especially in the downtown area; however, several very white community areas did not see a significant drop in ridership in the northern parts of the city, and several community areas with above 38% ethnic minority populations experienced large drops in ridership (particularly high ethnic-minority community areas near downtown). Displaying less of a distinct heterogeneous response by percent ethnic minority than by income. We use the following American Community Survey designations to define ethnic minority in this analysis: “Hispanic or Latino (of any race); Black and African American, Not Hispanic or Latino; American Indian and Alaska Native, Not Hispanic or Latino; Asian, Not Hispanic or Latino; Native Hawaiian and Other Pacific Islander, Not Hispanic or Latino; Two or More Races, Not Hispanic or Latino; Other Races, Not Hispanic or Latino”.

Appendix G: Regression for 2019: checking parallel trends

Table 6 shows the regression estimates, robust standard errors (clustered at the community area), and the significance for regression model 2 re-run for the same dates in 2019 with the artificial stay-at-home order as March 18th 2019. We see no significance in the interaction term or in the difference before and after the stay-at-home order date for the previous year. This indicates that the effects measured in the second regression in Table 2 are not due to seasonality or the date specifically, but due to the stay-at-home order and general pandemic fear leading people to avoid unnecessary trips.

Author contributions LH led the study design, data acquisition, analysis, interpretation of results, and manuscript preparation. CH contributed to study acquisition and design, and manuscript editing. DN contributed to study acquisition and design, and manuscript editing.

Funding Open Access funding provided by Carnegie Mellon University. The authors acknowledge support from the Block Center for Technology and Society at Carnegie Mellon University and the National Science Foundation Graduate Research Fellowship under Grant No. DGE1745016 & DGE2140739. This research was also supported by the US DOT Grant No. 69A3551747111 through Mobility21, a University Transportation Center, with the goal of improving the mobility of goods and services and the Scott Institute at Carnegie Mellon University.

Data and code availability The repository for this project can be found at https://github.com/Lilyhanig. The data can be found at https://data.cityofchicago.org/Transportation/Transportation-Network-Providers-Trips/m6dm-c72p

Declarations

Conflict of interest The authors have no competing interests to declare.

Ethical approval Not applicable.

References

Barajas, J.M., Brown, A.: Not minding the gap: Does ride-hailing serve transit deserts? J. Transp. Geogr. **90**, 102918 (2021). https://doi.org/10.1016/j.jtrangeo.2020.102918

Brodeur, A., et al.: Stay-at-home orders, social distancing, and trust. J. Popul. Econ. (2021). https://doi.org/10.1007/s00148-021-00848-z

Brough, R., et al.: Understanding socioeconomic disparities in travel behavior during the COVID-19 pandemic. Reg. Sci. (2021). https://doi.org/10.1111/jors.12527

Brown, A., Williams, R.: Equity Implications of Ride-Hail Travel during COVID-19 in California. Transport. Res. Record. (2021). https://doi.org/10.1177/03611981211037246

Callaway B. et al.: Difference-in-differences with a continuous treatment. (2021). doi: https://doi.org/10.48550/ARXIV.2107.02637

CDC/ATSDR SVI Data and Documentation Download (2021). url: https://www.atsdr.cdc.gov/placeandhealth/svi/data_documentation_download.html.

City of Chicago.: Transportation Network Providers Trips (2020a). url: https://data.cityofchicago.org/Transportation/Transportation-Network-Providers-Trips.

City of Chicago (2020b). COVID-19 Dashboard. url: https://www.chicago.gov/city/en/sites/covid-19/home/latest-data/2020b-03-23.html.

City of Chicago (2020c). Transportation Network Providers and Congestion in the City of Chicago. url: https://www.chicago.gov/city/en/depts/bacp/supp_info/transportation-network-providers.html

City of Chicago.: COVID-19 orders (2022). url: https://www.chicago.gov/city/en/sites/covid-19/home/health-orders.html.

Diao, M., et al.: Impacts of transportation network companies on urban mobility. Nat. Sustain. **4**, 494–500 (2021). https://doi.org/10.1038/s41893-020-00678-z

Erhardt, G.D., et al.: Do transportation network companies increase or decrease transit ridership? Empirical evidence from San Francisco. Transportation **49**, 313–342 (2021). https://doi.org/10.1007/s11116-021-10178-4

Goldstein M.: (2020). url: https://www.forbes.com/sites/michaelgoldstein/2020/07/27/what-is-the-future-for-uber-and-lyft-after-the-pandemic/?sh=5c6b04593bc8
Grahn, R., et al.: Socioeconomic and usage characteristics of transportation network company (TNC) riders. Transportation **47**, 3047–3067 (2020). https://doi.org/10.1007/s11116-019-09989-3
Hall, J.D., et al.: Is uber a substitute or complement for public transit ? J. Urban Econ. **1**, 1–5 (2018). https://doi.org/10.1016/j.jue.2018.09.003
Hanig, L., et al.: COVID-19 public transit precautions: Trade-offs between risk reduction and costs. Transport. Res. Interdiscipl. Perspect. (2023). https://doi.org/10.1016/j.trip.2023.100762
Hook, H., et al.: Does undirected travel compensate for reduced directed travel during lockdown ? Transport. Lett. **13**(56), 414–420 (2021). https://doi.org/10.1080/19427867.2021.1892935
Iio, K., et al.: COVID-19 and social distancing: Disparities in mobility adaptation between income groups. Transport. Res. Interdiscipl. Perspect. **10**, 100333 (2021). https://doi.org/10.1016/j.trip.2021.100333
Jin, S.T., et al.: Uber, public transit, and urban transportation equity: a case study in New York City. Prof. Geogr. **71**, 315–330 (2019). https://doi.org/10.1080/00330124.2018.1531038
Kantamneni, N.: The impact of the COVID-19 pandemic on marginalized populations in the United States: a research agenda. J. Voc. Res. (2020). https://doi.org/10.1016/j.jvb.2020.103439
Li, S., et al.: Regulating TNCs: Should Uber and Lyft set their own rules. Transport. Res. Part B: Methodol. **129**, 193–225 (2019). https://doi.org/10.1016/j.trb.2019.09.008
Liu, L., et al.: The impacts of COVID-19 pandemic on public transit demand in the United States. PLoS ONE (2020). https://doi.org/10.1371/journal.pone.0242476
National Center for Immunization and Respiratory Diseases (U.S.). Division of Viral Diseases.: In-terim guidance: Get your mass gatherings or large community events ready for coronavirus disease 2019 (COVID-19) (2020). url: https://stacks.cdc.gov/view/cdc/85893.
Nelson, E., Sadowsky, N.: Estimating the impact of ride-hailing app company entry on public transportation use in major US urban areas. J. Econ. Anal. Policy **1**, 1–5 (2019). https://doi.org/10.1515/bejeap-2018-0151
Ngo, N.S., et al.: The effects of ride-hailing services on bus ridership in a medium-sized urban area using micro-level data: evidence from the lane transit district. Transp. Policy (2021). https://doi.org/10.1016/j.tranpol.2021.02.012
Ozbilen, B., et al.: Perceived risk of infection while traveling during the COVID-19 pandemic: Insights from Columbus, OH. Transport. Res. Interdiscipl. Perspect. **1**, 1–5 (2021). https://doi.org/10.1016/j.trip.2021.100326
Pawar, D.S., et al.: Modelling work-and non-work-based trip patterns during transition to lockdown period of COVID-19 pandemic in India. Travel Behav. Soc. **24**, 46–56 (2021). https://doi.org/10.1016/j.tbs.2021.02.002
Qi, Y., et al.: Impacts of COVID-19 on public transit ridership. Int. J. Transport. Sci. Technol. **1**, 1–5 (2021). https://doi.org/10.1016/j.ijtst.2021.11.003
Rao L.: (2011). url: https://techcrunch.com/2011/09/22/uber-brings-its-disruptive-car-service-to-chicago/
Transit Center.: Tailoring Transit Service for Essential Workers is a Matter of Racial Justice (2020). url: https://transitcenter.org/tailoring-transit-service-for-essential-workers-is-a-matter-of-racial-justice/
Uber (2022). Mission Statement. url: https://www.uber.com/us/en/about/
US Census.: (2021). Chicago, Illinois Quickfacts. url: https://www.census.gov/quickfacts/chicagocityillinois
Ward, J.W., et al.: Effects of on-demand ridesourcing on vehicle ownership, fuel consumption, vehicle miles traveled, and emissions per capita in U.S. States. Transport. Res. Part C: Emerg. Technol. **108**, 289–301 (2019). https://doi.org/10.1016/j.trc.2019.07.026
Zipper D.: Post COVID transit needs a new metric for success (2021). url: https://www.bloomberg.com/news/articles/2021-02-16/post-covid-transit-needs-a-new-metric-for-success.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Lily Hanig is a Ph.D. student at Carnegie Mellon University in their Engineering and Public Policy department. She is in the Society, Policy, Infrastructure, Climate, and Energy (SPICE) Lab and the Future Mobility Systems Lab. She holds a B.E. in Energy Systems Engineering from Dartmouth College. Her research focuses on transportation resilience and decarbonization.

Destenie Nock is an Assistant Professor jointly appointed at the Department of Civil and Environmental Engineering and the Department of Engineering and Public Policy at Carnegie Mellon University. Her research aims to better understand how novel technologies in energy and transportation affect communities

and equality goals. She holds a B.S. in Electrical Engineering and Applied Mathematics from North Carolina A&T University, a M.S. in Leadership for Sustainable Development from Queen's University of Belfast, and a Ph.D. in Industrial Engineering and Operations Research from the University of Massachusetts Amherst.

Corey D. Harper is an Assistant Professor jointly appointed at the Department of Civil and Environmental Engineering and Heinz College at Carnegie Mellon University and Director of the Future Mobility Systems Lab. He holds a B.S. in Civil Engineering from Morgan State University and M.S. and Ph.D. degrees from Carnegie Mellon University from the Department of Civil and Environmental Engineering. His research seeks to understand better how emerging technologies in transportation affect infrastructure and communities.