Institute for Economic Studies, Keio University

Keio-IES Discussion Paper Series

Is Ride-sharing Good for Environment?

小西祥文、小野あかり

2024 年 6 月 6 日
DP2024-014
https://ies.keio.ac.jp/publications/23908/

Institute for Economic Studies, Keio University
2-15-45 Mita, Minato-ku, Tokyo 108-8345, Japan
ies-office@adst.keio.ac.jp
6 June, 2024

Is Ride-sharing Good for Environment?
小西祥文、小野あかり
IES Keio DP2024-014
2024 年 6 月 6 日
JEL Classification: L91, Q53, R4, R11
キーワード: Air pollution, congestion, commuting choice, staggered difference-in-differences, instrumental variable, ride-hailing, ride-sharing, transportation and environment

【要旨】

We estimate the causal effect of ride-hailing entry on transport-related air pollution in U.S. cities, using granular satellite-based NO_2 concentration data in the staggered difference-in-differences research design. Our empirical strategy accounts for treatment effect heterogeneity both within and across cities, coupled with two additional strategies to strengthen identification: using geography-based instruments and exploiting a sharp, unanticipated change in ride-hailing activity in Austin due to its rule change. We find robust evidence that ride-hailing tends to improve air quality in highly dense cities, but has no significant impact in cities with low and medium density. We also find evidence that the NO_2 reduction in highly dense cities is associated with a decrease in private car use and an increase in public transit use. Taken together, our findings suggest that the environmental effect of ride-hailing depends on the complementarity between ride-hailing and public transit: While ride-hailing may increase congestion by inducing deadheading or displacing of mass transit for parts of daily trips, it may still decrease overall air pollution if a combined use of ride-hailing with other transit displaces private car use more than such adverse behavior.

小西祥文
慶應義塾大学
〒108-8345 東京都港区三田2-15-45
ykonishi@econ.keio.ac.jp

小野あかり
慶應義塾大学
〒108-8345 東京都港区三田2-15-45
ono.akari@keio.jp

謝辞：The study was in part supported by the financial supports from Japan Society for the Promotion of Science (JSPS), Grant-in-aid for Scientific Research (Grant numbers: 18K01562, 21H00712, and 23K20599). Sho Kuroda, Takuya Shimamura, and Hazuki Yanagida provided valuable research assistance. We thank Michiko Namazu, a data scientist at Uber Inc., for sharing Uber's official launch dates in the U.S. We also thank Yuta Toyama, Toshi Arimura, and other seminar participants at Waseda University's Empirical Micreconomics Workshop for their comments.

Is Ride-sharing Good for Environment?

Yoshifumi Konishi*
Faculty of Economics
Keio University

and

Akari Ono
Faculty of Economics
Keio University

Draft: *June 2024*

*Corresponding author. Mailing address: 2-15-45 Mita, Minato-ku, Tokyo, 108-8345, Japan. E-mail: ykonishi@econ.keio.ac.jp. Phone: +81-3-5427-1725.

Acknowledgement: The study was in part supported by the financial supports from Japan Society for the Promotion of Science (JSPS), Grant-in-aid for Scientific Research (*Grant numbers: 18K01562, 21H00712,* and 23K20599). Sho Kuroda, Takuya Shimamura, and Hazuki Yanagida provided valuable research assistance. We thank Michiko Namazu, a data scientist at Uber Inc., for sharing Uber's official launch dates in the U.S. We also thank Yuta Toyama, Toshi Arimura, and other seminar participants at Waseda University's Empirical Micreconomics Workshop for their comments. This is a substantially improved version of the April 2021 draft.

Abstract: We estimate the causal effect of ride-hailing entry on transport-related air pollution in U.S. cities, using granular satellite-based NO_2 concentration data in the staggered difference-in-differences research design. Our empirical strategy accounts for treatment effect heterogeneity both *within* and *across* cities, coupled with two additional strategies to strengthen identification: using geography-based instruments and exploiting a sharp, unanticipated change in ride-hailing activity in Austin due to its rule change. We find robust evidence that ride-hailing tends to improve air quality in highly dense cities, but has no significant impact in cities with low and medium density. We also find evidence that the NO_2 reduction in highly dense cities is associated with a decrease in private car use and an increase in public transit use. Taken together, our findings suggest that the environmental effect of ride-hailing depends on the complementarity between ride-hailing and public transit: While ride-hailing may increase congestion by inducing deadheading or displacing of mass transit for parts of daily trips, it may still decrease overall air pollution if a combined use of ride-hailing with other transit displaces private car use more than such adverse behavior.

JEL Codes: L91, Q53, R4, R11

Key Words: Air pollution, congestion, commuting choice, staggered difference-in-differences, instrumental variable, ride-hailing, ride-sharing, transportation and environment

1. Introduction

> "*Ride-hailing is an attractive option for many travelers, and can increase mobility for households who lack a private vehicle. Yet in communities across the country, ride-hailing is increasing vehicle travel, climate pollution, and congestion.*" — *the Union of Concerned Scientists* (2020)

Over the last decade, Uber, Lyft, and other ride-hailing services have expanded rapidly, bringing innovations into the transport sector in numerous cities around the world. Economists have attempted to quantify the economic gains from this transport innovation for a number of important economic margins: Consumer's gain from Uber's congestion pricing (Cohen *et al.*, 2018; Castillo, 2023) and from reducing information asymmetry (Liu *et al.*, 2018) and drivers' gain from Uber's compensation scheme (Angrist *et al.*, 2021) and from flexible workstyle (Chen *et al.*, 2019). Economists have also refuted some of the criticisms against these ride-hailing companies, and instead, find: Uber's entry increased public transit riderships (Hall *et al.*, 2018), reduced driving under the influence, fatal accidents, arrests for assault and disorderly [Dills and Mulholland (2018), Anderson and Davis (2023)], and had no significant effects on taxi driver's labor supply although reducing their relative earnings by about 10 percent (Berger *et al.*, 2018). Against these economic benefits, however, a rising number of studies have also started to document the negative environmental impacts of ride-hailing services [Diao *et al.* (2021), Erhardt *et al.* (2019), Kong *et al.* (2020), Rayle *et al.* (2016), Tarduno (2021), Krishnamurthy and Ngo (2024)].

Ride-hailing is thought to increase air pollution and congestion primarily for two reasons — 'deadheading' (miles without a passenger between hired rides) and 'displacing' (miles that might have displaced mass transit or other low-emission travel modes). For example, the Union of Concerned Scientists (UCS) estimates that the former contributes 47% more emissions per trip while the latter adds 69% more per trip, based on the data from U.S. cities (UCS, 2020). Such claims often receive support from observational studies that document a surge in road traffic upon ride-hailing entry (e.g., T&E, 2019). While we agree with the general sentiment of these studies, they miss an important economic channel: Complementarity between ride-hailing and other transit modes. Ride-hailing customers may use public transit in combination with hired rides, and such a combined use may displace the use of private driving, not only for primary commuting but also for daily errands. Put differently, the UCS study compares hired rides with other transport modes that could have replaced the rides, but that isn't a valid counterfactual if the aforementioned channel exists. In this case, a valid counterfactual comparison would be, instead, to compare *a sequence of daily transport* *choices* in the presence of ride-hailing against those in its absence. The goal of our study is to establish credible evidence on the causal effect of ride-hailing on ambient air pollution, based on the empirical design better suited to make such a counterfactual comparison.

To do so, we start by the canonical two-way fixed effect (TWFE) regression with the staggered difference-in-differences (DD) design, exploiting the variation in ride-hailing entry over time and across cities. Our basic empirical strategy is to estimate this TWFE-DD regression on monthly observations of ambient air pollution on a panel of 348 MSAs during the 9-year period, 2010-2018. Thus, our causal inference relies on how we carefully design treatment-control structures on the study sample. This basic strategy itself is analogous to previous studies [Berger *et al.* (2018), Hall *et al.* (2018), Li *et al.* (2021), Ward *et al.* (2019), Kim and Sarmiento (2021)].

We, however, take three new approaches to strengthen the identification of the causal effect. First, we construct a measure of *de facto* ride-hailing entry into a MSA boundary, using both Google's keyword search trends for both Uber and Lyft. This measure of entry is more complete in coverage, is more accurate in both entry timing and location, and also helps us avoid the possibility of falsely refuting the critics' argument in favor of ours. We demonstrate these points more forcefully in **Section 4**.

Second, we use satellite-based nitrogen dioxides (NO_2) concentration data. We specifically avoid use of U.S. Environmental Protection Agency (EPA)'s monitoring-based data on other pollutants for the identifiability of ride-hailing's impacts on *transport-related* air pollution. Monitoring data may not be reliable for credible inference for various reasons: Monitoring sites are spatially unevenly located and are not necessarily located in high pollution areas (Fowlie *et al.,*2019); site locations change over time, making it difficult to compare data consistently over time; and there is an "unwatched pollution problem" in that local governments may have incentives to strategically locate monitoring sites or avoid recording high pollution episodes (Zou, 2021; Grainger and Schreiber, 2019). In contrast, we use the satellite-based NO_2 concentration data to calculate the ambient air pollution on spatially delineated subareas within each MSA. This allows us to compare, for example, urban-area NO_2 concentrations of ride-hailing entry city against that of no-entry city consistently over time.

Third, we take three carefully designed identification strategies so as not to rely on a single set of identifying assumptions for credible inference. The major identification threat to the staggered DD design is that ride-hailing entry may occur in cities exactly when and where residents' transport-related behavior is expected to change. This leads to the violation of the parallel trend in unobservables. Although we include a MSA-specific linear trend along with other fixed effects and time-varying controls, this may not completely eliminate all unobservables that are correlated with both the treatment and the outcome. Our first strategy is to mimic the idea of matched DD (Heckman *et al.*, 1997). We start by noting that entry dates are highly correlated with population density, which are also correlated with other MSA-level socioeconomic variables such as income, manufacturing employment, and the share of public transit commuters. Hence, we use MSA-level population density as a sufficient statistic for unobservable trends in transport-related behavioral change, and estimate the causal effect on each subsample of MSAs stratified by population density quintile. We then employ the de Chaisemartin-D'Haultfoeulle (dCDH, 2020; 2024) estimator to estimate heterogeneous dynamic treatment effects.[1] The second strategy exploits the sharp, unanticipated change in the supply of ride-hailing service in Austin, Texas, due to its rule change on fingerprint checks on ride-hailing drivers. We apply the synthetic control method (SCM) on the residualized outcomes to visualize the impacts of both the entry and the rule change on Austin's air quality, using this incidence as a quasi-experiment. Our last strategy applies an instrumental variable (IV) method to the TWFE-DD regression. Our instruments are *reported* entry dates, taken from Hall *et al.* (2018), interacted with the geography-based instruments, which are widely used in the empirical economic geography literature [Baum-Snow (2007), Duranton and Turner (2011; 2012), Faber (2014), Redding and Turner (2015)]. The basic idea here is that the geography-based instruments create 'hypothetical highway routes', which predict the current routes (hence, the current economic size of the cities) well, yet are not correlated with contemporaneous economic shocks after controlling for observables today.

Using these approaches, we find robust evidence that ride-hailing entry tends to decrease ambient NO_2 concentrations (in terms of both monthly mean and maximum) in MSAs with high population density, particularly in their urban areas. We also find no evidence of ride-hailing entry leading to an increase in ambient NO_2 concentrations in low- and medium-density MSAs. We also confirm these findings in the event study of dynamic treatment effects using a version of the dCDH estimator. The signs of the estimated impacts are largely consistent across different identification/estimation strategies, although their magnitudes vary. In particular, our DD-IV estimation leads to unreasonably large estimates in lower density MSAs. We suspect that given the form of the DD-IV estimand, the lack of reported entry dates for these MSAs overly inflate the estimates. Nonetheless, the estimates for the urban areas of MSAs in the highest density quartiles are always negative, have relatively small standard errors, and range from -0.034 log points (TWFE-DD) to -0.104 log points (DD-

[1] Alternatively, we may use the Callaway-Sant'Anna (CS) estimator, which is theoretically identical with the dCDH estimator in the absence of covariates. As we explain in **Section 5**, we absorb the influence of all time-varying covariates using the imputation method (Catenao *et al.*, 2023). Hence, we should obtain the similar estimates using either estimator.

IV). Assuming a linear relationship between satellite-based and EPA monitoring records, these impacts translate into the reductions of monthly mean NO_2 concentrations by roughly 0.53-1.62 ppb.

As a supplement to our main analysis, we also use annual household-level data on commuting modes to work from the American Community Survey (ACS) to further explore the economic mechanism underlying our main results. For this, we use commuting mode indicators as the outcome variables, and apply all three identification strategies discussed above. The results are quite supportive to our argument. We find that the air-pollution impacts of ride-hailing is significantly associated with changes in commuting patterns. In the highest density MSAs, where ride-hailing entry is estimated to decrease NO_2 concentrations, ride-hailing entry is also significantly associated with a decrease in private car commuters, and increases in public transit commuters and in the other commuting modes. Interestingly, we also find that the sharp decrease in the supply of ride-hailing in Austin due to its rule change is estimated to reduce, rather than increase, the ambient NO_2 pollution and the share of private car commuters while increasing the share of public transit and other commuting modes. Taken together, these results support our argument that the environmental effect of ride-hailing depends critically on the degree of complementarity (or substitutability) between ride-hailing and other transit modes. In highly dense cities, such a complementarity is high so that a combined use of ride-hailing with mass transit can reduce private car use whereas in lower dense cities, the complementarity is weak so that ride-hailing tends to steer people away from mass transit. These results are in sharp contrast to critics' views cited above, but are indeed consistent with Hall *et al.* (2018), who find complementarity between ride-hailing and other transport modes.

Our work complements several vibrant areas of research: (a) empirical studies that estimate the causal effects of ride hailing on various economic outcomes [Anderson and Davis (2023), Angrist *et al.* (2021), Berger *et al.* (2018), Chen *et al.* (2019), Cohen *et al.* (2018), Dills and Mulholland (2018), Hall *et al.* (2018), Liu *et al.* (2018), Tarduno (2021), Krishnamurthy and Ngo (2024)], (b) a large body of literature that examines pollution- or congestion-relief effect of public transportation infrastructures [Chen-Whalley (2012), Li *et al.* (2019), Gendron-Carrier *et al.* (2022), Gu *et al.* (2021), other papers cited in Anas and Lindsey (2011)], and (c) the economic studies that structurally investigate the general equilibrium impacts of ride-hailing into the taxi industry [Buchholz (2023), Fréchette *et al.* (2019), Hall *et al.* (2020), Rosaia (2023)].

Of these, our work is most closely related to Krishnamurthy and Ngo (2024) and Kim and Sarmiento (2021). Krishnamurthy and Ngo (2024) use hourly freeway traffic and daily air pollution data from California and apply a DD design exploiting the staggered rollouts of ride-hailing entry at the country-level. They find that ride-hailing entry reduces weekday freeway congestion and $PM_{2.5}$ concentrations in the average county entered, although congestion and $PM_{2.5}$ concentrations are increased during the evening rush hour and in the most populated counties. Our study nicely complements theirs in that our findings are quite consistent, yet our scope is quite different from theirs — we use more aggregate data for virtually all MSAs in the U.S. while they use more disaggregate data only for California. Kim and Sarmiento (2021) use the Callaway-Sant'Anna estimator in an empirical design similar to ours and find that Uber's entry is estimated to improve air quality (as measured in the Air Quality Index and ground-level ozone), particularly during the summer when bad air quality episodes are expected. Our study complements and further strengthens their findings in four regards. First, while they rely mostly on *state-level reported* entry dates of Uber, we construct MSA-level de facto entry dates from the Google Trends Index for both Uber and Lyft. We show in **Section 4** that our entry dates are likely to be more complete in its coverage, more accurate in both location and timing, and correctly capture the ride-hailing activity. Second, while they rely on EPA's monitoring data and use county-level observations, we rely on satellite-based data and urban areas of MSAs as study units. Third, while they rely on one identification strategy (staggered DD with various specifications, controls, and robustness checks), we take three alternative identification strategies and find consistent results. Fourth, while they focus only on ambient air quality as outcomes, we also explore the economic mechanism underlying the main results. We find the estimated impacts of ride-hailing entry on commuting modes are indeed consistent with the estimated impacts on ambient air quality.

Lastly, our analysis gives a clear answer to an ongoing debate among policy practitioners and scientists (i.e., "Does ride-hailing decrease or increase air pollution and congestion?) in ways that can embrace and reconcile seemingly conflicting empirical findings in the literature. On one hand, studies based on disaggregate data tend to find that ride-hailing increases congestion in *particular* segments of transport demand (e.g., rush hours, urban freeways, populated counties) while decreasing congestion in *other* segments (e.g., non-rush ours, non-urban freeways, less populated counties) [Erhardt *et al.* (2019), Tarduno (2021), and Krishnamurthy and Ngo (2024)]. On the other hand, studies that rely on more aggregate data such as MSA-level or monthly-level data tend to find that ride-hailing increases use of mass transit (Hall *et al.,* 2018), decrease congestion (Li *et al.*, 2021), decrease ambient air pollution (Kim and Sarmiento, 2021), and vehicle emissions (Ward *et al.*, 2019). Our results suggest that the key to reconciling these mixed findings is the heterogenous effects of ride-hailing on various segments of transport demand. That is, ride-hailing may decrease air pollution by encouraging a combined use of ride-hailing and mass transit in cities that have dense public transit networks, yet may increase congestion in some part of the cities, due to deadheading by ride providers or use of hired rides in place of mass transit for *part of the daily trips (but not the entire sequence of daily trips).* Our results is also consistent with Agrawal and Zhao (2023), who argue, based on a simulation-based study using the monocentric city model, that subsidizing ride-hailing services as a “last-mile” provider can be welfare-increasing and whether ride-hailing and public transit are substitutes or complements is a policy choice. Thus, our manuscript provides an important insight into an important policy debate that seeks to strike a balance between increasing mobility and fighting air pollution in cities around the world.

2. Background and Motivation

2.A. Ride-hailing Service and Its Environmental Concerns

Uber Technologies Inc. ("Uber") started as a developer of a smartphone application that would make ride-hailing as simple as "tapping a button". Uber launched its first ride-hailing service in San Francisco in July, 2010 and in New York in May, 2011. During the initial phase, Uber mainly operated the ride-hailing platform for expensive limousines ("black car"), but later introduced a more affordable service, UberX. UberX was first launched in San Francisco in January, 2013, quickly became Uber’s standard ride-hailing service throughout the U.S., and is often seen as a direct competitor against the traditional yellow cab service. Lyft Inc. ("Lyft"), on the other hand, started as a long-distance ride-sharing service between college campuses in 2007, with a brand name Zimride. Its first short-distance ride-sharing service appeared in San Francisco in August, 2012 as a complementary service of Zimride. In 2013, the company changed its name from Zimride to Lyft and sold it to Enterprise Holdings. Uber and Lyft have entered roughly 80% of U.S. cities and stayed active there. Today, Uber accounts for roughly 70% of the ride-hailing service sales in the U.S., with Lyft holding the remaining 30%, and users are royal to their service providers — most customers use only one service and rarely switch services (Second Measure, 2021).

Ride-hailing services are known to provide a number of economic benefits: ease of access to transit, flexible workstyle, increased employment opportunities for the poor, reducing information asymmetry and mismatch in the taxi market, and promoting smartphone-based innovations in other areas of the economy. Against these benefits, however, they are often criticized for the downsides of their business model. One of the important controversies is whether ride-hailing services decrease or increase congestion and road traffic, particularly in dense urban metro areas. Increased congestion and road traffic, if true, are a cause of serious concerns from an environmental perspective. Internal combustion of fossil fuels by vehicles is the leading source of harmful air pollutants such as CO and NO_2. Such pollutants are known to increase the risk of stroke, heart disease, lung cancer, and chronic and acute respiratory diseases. There is a large body of literature that establishes a positive relationship between increased congestion/traffic and air pollution/carbon emissions from on-road vehicles [See Anas and Lindsey (2011) for a nice review on this issue]. Hence, the congestion problem could also be directly linked to climate and pubic health concerns. Recent reports from two high-profile organizations tout for such concerns [T&E (2019); UCS (2020)].

There are two competing views on the environmental effect of ride-hailing. On one hand, proponents of Uber and Lyft (including the companies themselves) argue that ride-hailing apps provide easy access to shared mobility on demand, allowing commuters to rely less on private car ownership and more on public transit and other transport modes such as bicycling and walking. On the other hand, critics argue that ride-hailing can add road traffic either because of 'deadheading' (driving without a passenger between hired rides) or because ride-hailing increases reliance on hired rides rather than on public transit and other transit modes. Previous studies indeed find mixed results on this issue. Studies that rely on interviews, surveys, and micro-level trip data in specific cities [Erhardt *et al.* (2019), Kong *et al.* (2020), Rayle *et al.* (2016); UCS (2020)] tend to find results in support of the critics' arguments while studies that rely on quasi-experimental research design [Hall *et al.* (2018); Li *et al.* (2021); Ward *et al.* (2019)] tend to refute the critics' arguments.[2]

One point we wish to clarify in this paper is that these mixed findings arise mostly because they make empirically quite different counterfactual comparisons. Studies that find in support of the critics' arguments compare hired rides with other transport modes that could have potentially replaced the rides. For example, in the UCS study, a hired ride with an average level of deadheading is compared against other modes of travel an average commuter might opt for in the absence of ride-hailing services. However, in our view, that is not a fair or valid counterfactual. Consumers make a sequence of daily transport choices, and hence, use ride-hailing in combination with other transport modes (We discuss this point more fully in the next section). A valid counterfactual in this case, instead, is to compare a sequence of transport choices over a course of the day in the presence of ride-hailing against those in its absence. In other words, we need a counterfactual comparison that accounts for behavioral

[2] An exception is Diao *et al.* (2021), who uses a difference-in-differences research design analogous to Hall *et al.* or Li *et al.*, but finds, in contrast, that ride-hailing entry increases congestion and decreases public transit ridership. In our view, however, Diao *et al.* fails to deliver credible results because they fail to control for MSA-specific time trends and instead use mostly endogenous controls and highly suspicious IVs along with a prohibited second-stage regression using the predicted probability from the first-stage logit regression.

changes in equilibrium commuting patterns in a city. We discuss this point more forcibly in the **next section**.

2.B. Why May Ride-hailing Decrease or Increase Transport Emissions?

U.S. cities have experienced substantial suburbanization over the last half century. Central city population declined by 17% whereas total MSA-level population increased by 72% between 1950 and 1990 for large MSAs, due primarily to the rapid development of limited access highways over this period (Baum-Snow, 2007). As a result of this suburbanization, a majority of MSA residents make either suburb-to-central city or suburb-to-suburb commute for their work. About 63% of MSA commuters make such trips and about 87% use private cars for daily commuting in 2000 (U.S. Census Bureau, Journey to Work). Thus, the impact of ride-hailing entry on transport-related emissions depend on how it affects commuters' transport choices in such a suburbanized city.

In this regard, a recent empirical finding by Hall *et al.* (2018) — that ride-hailing services are a complement to public transportation, particularly in large cities — gives us an important insight. In large cities where commuters have access to sufficiently dense public transit networks, ride-hailing can complement public transit in a variety of ways. For example, public transit stations may be far from commuters' homes but may be close to their workplace. Ride-hailing can connect such commuters to their nearest transit, potentially allowing them to switch from private driving to public transit. For another example, commuters may run a variety of chores while at work. Commuters may opt to drive their own car to work, not just for convenience of commuting but for such anticipated chores. Low-cost, easy-to-hail ride services may alleviate this latter need for driving their own cars. This line of reasoning suggests that ride-hailing can potentially reduce private driving *more than* simply replaces it — the ride-hailing service is used in place of private driving, but the use of public transit (or ride-hailing) for other parts of travel that comes with it can replace private driving as well.

We demonstrate this point using Chicago as an illustrative example. **Figure 1** shows major train and bus routes in the urban and suburban areas of Chicago, along with a stylized diagram intended to conceptualize a certain segment of the city. As with many cities in the U.S., Chicago has a radial network of public transit routes that are highly concentrated around the central business district (CBD) and extend radially outward from the CBD to the suburban areas. As a result, many suburban areas have sparse transit networks that cannot directly connect their residents to either the CBD or other employment centers.

Imagine a commuter who lives in a suburb and commute to a city's employment center for work. Often, the nearest station is not located within a walking distance from/to commuter's home or workplace. Hence, the commuter may need to take another mass transit to the nearest station. Given this inconvenience, the commuter may opt to drive her car to work in the absence of low-cost, easy-to-hail ride services. This driving distance is denoted by s_c in the diagram. In addition, the commuter may use her car to run a few 'daily chores' while at work: attending meetings at other places, buying things at stores, going to restaurants, etc. For simplicity, the driving distance is denoted by a radius s_e from her workplace. In the presence of convenient ride-hailing services, however, the commuter could hire rides to the nearest station, either from her home or from her workplace, and take the mass transit instead. The driving distances by hired rides in this case are denoted r_h and r_w, respectively. Furthermore, such a commuter may also use hired rides to replace the driving distance for daily chores s_e. Assuming that the mass transit operation stays the same, the total amount of air pollution depends only on the total driving distance, which would change from $s_c + s_e$ to $r_h + r_w + s_e$ if a complementarity between ride-hailing and public transit exists. In this case, ride-hailing would decrease the total driving distance if $s_c > r_h + r_w$, yet increase the congestion in the urban area because $s_e < r_w + s_e$. Of course, if such complementarity does not exist or is not strong enough, hired rides might simply displace mass transit commuting m. **Figure A1** in the **Appendix** provides some anecdotal evidence in support of such complementary use of ride-hailing services. The figure displays the observed geographic distribution of ride-hailing trips during the weekday rush hours in Chicago, demonstrating that roughly 75% of the rush-hour trips from the suburban community areas do not travel the full distance to the CBD.

It is, therefore, largely an empirical question to what extent such complementarity exists, and whether it is strong enough to outweigh the increase in driving activity by ride-service providers. In addition, these discussions suggest that the impact of ride-hailing can be highly heterogeneous across cities because the degree of complementarity (or substitutability) between ride-hailing and mass transit is likely to depend on the density of public transit.

3. Estimation and Identification Strategy

3.A. Overview

Our empirical strategy employs a staggered difference-in-differences (DD) research design, building on earlier empirical studies [Berger *et al.* (2018); Hall *et al.* (2018); Li *et al.* (2021)].

For all analyses, we use metropolitan statistical areas (MSAs) as our study units. Our study covers the period 2010-2018 and uses the 2009 core-based statistical area (CBSA) boundaries for all years' observations. This coverage ensures that each treated MSA has sufficient leads and lags before and after its ride-hailing entry, and that the period does not overlap with Uber Eats activity.

As a basis of our analysis, we start with the standard two-way fixed effect (TWFE) specification of the staggered DD regression, following Hall *et al.* (2018) and Berger *et al.* (2018). For an outcome variable y_{ctm} in year t and month m of city (MSA) c, we specify our TWFE regression as:

$$y^j_{ctm} = \alpha_c + \lambda_{tm} + \sum_s \delta_s D_{ctm} + X'_{ctm}\gamma + \theta_c(t) + \epsilon_{ctm}, \tag{1}$$

where D_{ctm} is our treatment variable and equals 1 if ride-hailing service enters/exists in MSA c and 0 otherwise in period tm, δ_s is the heterogenous treatment effect parameter corresponding to s-th quartiles of MSA-level population density with $s \in \mathcal{S} \equiv \{1, 2, 3, 4\}$, X_{ctm} is a vector of (exogenous) time-varying covariates, α_c and λ_{tm} are MSA and year/month fixed effects, and $\theta_c(t)$ is a MSA-specific linear time trend. Our main outcome variables of interest are ambient levels of nitrogen oxides (NO_2) concentrations in logged terms. **Section 4** justifies the use of NO_2 concentrations as the most appropriate measure of transport-related air pollution for our purpose. We use two measures of our outcome: the MSA-level means (over girded cells) of (1) monthly average NO_2 concentrations and (2) monthly maximum NO_2 concentrations. Time-varying controls include temperature, wind speed, their polynomials, gasoline prices, and non-attainment status for 1997 CAAA standards for PM2.5 and O_3. Temperature, wind speed, and gasoline prices are aggregated at the state level to avoid substantial data attrition as well as confounding with our treatment.[3] We carefully choose our time-varying controls since inclusion of endogenous time-varying controls is known to cause severe bias in the estimates. The population density quintile is based on the 2010 data to ensure that it is the pre-treatment status.

There are several important empirical challenges in estimating eq. (1), and we address each of them as follows.

First, earlier studies use *reported* Uber/Lyft entry dates, making use of information drawn from local newspapers, official blogs/websites, and social networking services. These

[3] We use temperature and wind data from the monitoring records available at the EPA's AirData. As we shall discuss in **Section 4**, monitoring data are available only sparsely. Hence, a non-negligible share of our sample would need to be dropped if we were to aggregate these data at the MSA level. Furthermore, we would like only the supply-side variation in gasoline prices, excluding the influence of MSA-level demand-side factors for gasoline consumption.

reported dates may have incomplete coverage, potentially multiple dates of entry, and some reporting errors. More importantly, in some cities, reported entry occurs well before the ride-hailing service achieves sufficient market penetration. This endangers the risk of false negative by construction — i.e., Uber/Lyft would not increase transport-related air pollution until a sufficiently large number of commuters respond to it. To overcome this problem, we construct a measure of *de facto* entry, constructed from the Google Trends Index. In the **next section**, we define our measure of *de facto* entry and discuss its empirical properties more fully.

Second, recent advances in the environmental economics literature indicate that the use of EPA's monitoring data may lead to substantially biased statistical inferences. EPA's monitoring stations are spatially unevenly located, and many of them are discontinued or change locations over time (Fowlie *et al.*, 2019). Furthermore, there is an "unwatched pollution problem" in that local governments may strategically choose monitoring sites or avoid recording high air pollution episodes (Zou, 2021; Granger-Schreiber, 2019). These characteristics make it very hard to credibly compare pollution data over time and across space. Instead, we make use of the National Aeronautics and Space Administration (NASA)'s satellite-based NO_2 concentration data, which are available at relatively high resolutions ($0.25° \times 0.25°$ girded cells) consistently throughout our study period. We discuss our NO_2 data more thoroughly in **Section 4**.

Third, although NO_2 has a known distance-decay relationship (Cape *et al.*, 2004; EPA, 2008), it may still travel far distance and affect ambient concentration levels several miles away from its emissions sources (Su *et al.*, 2009).[4] This implies that the violation of STUVA (or no-spillover) condition may occur in several ways, particularly between neighboring MSAs. To avoid such spillover effects, we delineate urban, suburban, and non-urban boundaries within each MSA, calculate the monthly NO_2 concentration statistics for each of these subareas $j \in \{\text{urban, suburban, non-urban}\}$, and run a separate regression for each j in eq. (1). This allows us to compare, for example, urban-area NO_2 concentration of entry city against urban-area NO_2 of no-entry city. Because urban areas are quite far apart from each other, this minimizes the risk of STUVA violation. Furthermore, our discussion in **Section 2** suggests that ride-hailing entry may have different air-pollution impacts on different subareas of cities. Since we expect some spatial spillovers across subareas of cities, how the

[4]It is known that the effect of emissions from on-road vehicles on ambient NO_2 concentrations declines quickly with distance — 90% of the decline occurs within just a 10-meter distance and the return to the baseline concentration levels occurs between 200 and 500 meters [see EPA (2008) and papers cited therein]. However, recent studies have shown that the influence of NO_2 emissions may extend beyond this conventional distance range through complex reactive/mixing processes with background pollutants in the atmosphere. Applying the land-use regression models in Los Angels, Sue *et al.* (2009), for example, found that the spatial extent of influence can reach as far as 5-20 kilometers from the emissions source.

treatment effects differ over subareas of cities is also an important empirical question.

Fourth, the major threat to identification in eq. (1) comes from the violation of parallel trends in unobservables. As discussed in Hall *et al.* (2018), ride-hailing services mostly enter cities in the rank order of population size (or population density). However, even if ride-hailing entry is mostly exogenous to residents' transport choices, the parallel-trend assumption may still be violated because Uber/Lyft happen to enter cities exactly where and when transport-related air pollution is expected to decline (say, due to preferences/public efforts toward more eco-friendly transportation behavior) or is expected to rise (say, due to growth of economic activities). Indeed, we show in the **Appendix** that entry dates are highly correlated with population density, the share of manufacturing employment, and the share of workers who use public transit for daily commuting. To overcome this challenge, we employ three identification strategies. In this section, we discuss our first strategy below, leaving the other two in **Sections 6 and 7**, respectively.

Our first strategy mimics the idea of matched difference-in-differences approach (Heckman *et al.*, 1997). Ideally, we would compare the outcomes of cities with similar air pollution trends in the absence of Uber/Lyft entry. If we have a large sample of cities with large variation in timing of Uber/Lyft entry, we would match cities based on all available observables. We, however, have only 348 MSAs with little variation in entry timing. Thus, matching on even a few observables can quickly exhaust observations that can be used as control units because Uber/Lyft enter similar cities at similar timings. This makes it infeasible to directly apply the matched DD method. Instead, we rely on the (arguably heuristic) argument that population density is likely a sufficient statistic for unobserved trends. Specifically, our first approach makes the following identifying assumption:

A1. *Conditional on exogenous (time-varying) covariates* $\Omega_c \equiv \{\alpha_c, \lambda_{tm}, X_{ctm},\ \theta_c(t)\}$, *parallel trends in unobservables hold for all MSAs in the same population density cohort:*

$$E[\epsilon_{c,\tau} - \epsilon_{c,\tau-1} | D_{c\tau} = 1, \Omega_c] = E[\epsilon_{c,\tau} - \epsilon_{c,\tau-1} | D_{c\tau} = 0, \Omega_c],$$

for each city c *within each population density quartile* s *for all periods* τ.

Note that in the staggered DD setup, parallel trends (in unobservables) need to hold for each timing group: i.e., unobservables for cities treated in g-th period would move in the same way as for all cities not treated in that period [Goodman-Bacon (2021), Sun and Abraham (2021), Callaway-Sant'Anna (2021)]. The assumption **A1** helps relax this assumption by requiring the parallel trends for each timing group to hold only within the same density cohort. We take two alternative estimation strategies under this assumption.

The first is to estimate the treatment effect parameter separately on each density-quintile subsample. This approach is taken in our event-study estimation in **Subsection 5.B** as well as our instrumental variable approach in **Subsection 7.B**. The second is to fully interact the population density quintile dummies with the treatment variable. The latter is the approach we take with our TWFE specification in eq. (1).

The specification in (1) also helps our identification in two other regards. As discussed in Hall *et al.* (2018), ride-hailing services mostly entered cities in the rank order of population size (or, more precisely, population density as shown in the **Appendix B**). Because we expect the treatment timing to vary by population density, our specification is a parsimonious way to interact the timing group dummies ($\approx$ the population density quintile dummies) with the treatment variable. This is known to remove bias arising from the possible correlation between treatment effect heterogeneity and the timing of entry [Sun and Abraham (2021), Wooldridge (2021)]. Furthermore, our specification also helps us explicitly account for heterogeneous treatment effects of ride-hailing entry to vary by (pre-treatment) population density. This is important because we expect the complementarity between ride-hailing service and public transit to be greater in cities with density (both public transit density as well as proximity to each other are important). Słoczyński (2020) has shown that when such heterogeneity exists, the best way to identify the true treatment effect is to interact the treatment with the sources of heterogeneity rather than to include them as controls.

3.B. Event Study

It is now increasingly common to estimate dynamic treatment effects, using event-study specifications. Recently, a series of studies have documented that severe biases may arise in estimating TWFE event-study regressions in the staggered DD setup, particularly when the effects are both dynamic and heterogeneous [de Chaisemartin-D'Haultfoeulle (2020, 2024) Goodman-Bacon (2021), Sun and Abraham (2021), Callaway-Sant'Anna (2021)]. In the worst case, the bias may be so severe that the TWFE estimates may have the signs that are opposite to the true parameters. Several alternative estimators are proposed that can properly address the problem. Among others, de Chaisemartin-D'Haultfoeulle (dCDH) and Callaway-Sant'Anna (CS) estimators are probably the most robust to misspecifications (and probably the most widely used). The problem with the CS estimator, however, is that it does not allow for time-varying covariates. In our context, air pollution concentration is highly seasonal and is affected by time-varying climatic conditions such as wind and temperature. Hence, failing to control for the effect of such time-varying confounds is likely to give us biased estimates. Hence, to address this issue, we exploit the imputation approach suggested by by

Gardner (2021) and Catenao *et al.* (2023): we can use pre-treatment observations to purge out the effect of (exogenous) time-varying confounders. Specifically, we take the following steps:

1. Estimate the following using only pre-treatment data:

$$y^j_{ctm} = \alpha_c + \lambda_{tm} + X'_{ctm}\gamma + \theta_c(t) + \epsilon_{ctm} \tag{2}$$

2. Estimate "residuals" using all data (incl. post-treatment):

$$\hat{\epsilon}^j_{ctm} = y^j_{ctm} - (\hat{\alpha}_c + \hat{\lambda}_{tm} + X'_{ctm}\hat{\gamma} + \hat{\theta}_c(t))$$

3. Apply the following dCDH/CS estimator to the estimated residuals, separately for each subarea j and each density-cohort subsample s (not pooling all MSAs): For MSAs that receive treatment in g-th period, estimate the average treatment effect on the treated (ATT) for τ-th period as[5]

$$ATT^{j,s}(g,\tau) = E[\hat{\epsilon}^j_{c,\tau} - \hat{\epsilon}^j_{c,g-1}|G_g = 1] - E[\hat{\epsilon}^j_{c,\tau} - \hat{\epsilon}^j_{c,g-1}|C = 1] \tag{3}$$

   where G_g is an indicator of g-th timing group (i.e., equals 1 if MSAs receive treatment in g-th period for the first time and 0 otherwise), and C is an indicator of not-yet-treated MSAs as of τ-th period (or never treated MSAs if such MSAs exist).

There are several reasons why this approach using steps 1-3 may yield more consistent estimates of ATTs than the TWFE regression (1). First and the foremost, the dCDH/CS estimator removes the bias that arises from heterogeneous dynamic treatment effects. The primary reason for such a bias is that the TWFE regression uses 'already treated' observations as effective control units when estimating the impact of entry on 'later treated' units. Although our specification mitigates this problem by including interactions with density cohort dummies, the dCDH/CS estimator removes the bias all together by specifically avoiding use of such units. Second, by step 1, we remove the potential bias that may arise due to time-varying covariates. Time-varying covariates are known to produce biased estimates if

[5] This is a version of the CS estimator without covariate adjustment, which is essentially the same as the dCDH estimator in the absence of switchers. The original CS estimator adjusts weights for different units using generalized propensity scores estimated on pre-treatment time-invariant covariates. As discussed in the previous subsection, we instead use pre-treatment population density as a sufficient statistic and estimate the dCDH/CS estimator separately on each density-quintile subsample.

they are correlated with treatment assignment. We avoid the bias by estimating the parameters using pre-treatment observations. Third, this approach explicitly makes use of the assumption **A1,** and estimate ATTs using only observations in the same density cohort. Hence, our approach compares only 'similar units' in terms of the pre-treatment population density.[6]

One potential drawback of this estimator, however, is that we use only 'relevant' observations that provide valid comparisons. Thus, this estimator tends to produce larger standard errors than the TWFE regression when there are a small number of comparable units. Hence, we should keep in mind that there is a trade-off between consistency and efficiency in interpreting the estimates from the two estimators.

4. Data

A. Ride-hailing Entry: Given our goal of making the counterfactual comparison discussed in **Section 2**, the chronology of the ride-hailing industry makes it difficult to define 'entry' of ride-hailing service in a city for a number of reasons. First, earlier studies use *reported* Uber/Lyft entry dates, making use of information drawn from local newspapers, official blogs/websites, and social networking services. These reported dates are often incomplete and sometimes inaccurate. Second, Uber and Lyft enter each city in different timings, with varying levels of market presence. Yet, presence of either of the two companies may be sufficient to induce changes in commuting patterns. Third, the official dates of entry may be an imprecise measure of a ride service's penetration into the city's market. We need a measure of *de facto* entry that embodies the sufficient market penetration. This is particularly important in our study context. Changes in transport choice would not have occurred on a large scale, say, in San Francisco immediately after Uber's entry in 2010 or UberX's entry in 2013. Defining this date as ride-hailing entry would falsely refute critics' argument because Uber's entry probably had no immediate effect on ambient air pollution almost by definition. Put differently, we should see the negative effect of ride-hailing (if critics are right) only after ride-hailing sufficiently penetrated the city's taxi market.

To overcome these challenges, we use Google Trends data to construct a measure of ride-hailing entry, building upon Hall *et al.* (2018) argument that normalized Google Trends closely coincide with monthly trends in active Uber drivers in their sample of U.S. cities. **Figure 2** reinforces Hall *et al.*'s as well as our arguments. In **Figure 2-(a)**, we plot normalized Google Trends indices for Uber and Lyft in San Francisco. Search activities were the

[6] We, however, do not use a generalized propensity score to condition on other pre-treatment covariates X unlike the original CS estimator. As discussed above, matching on pre-treatment covariates was infeasible due to the small sample size with little variation in entry timing.

highest in 2017 in San Francisco for both Uber and Lyft. Relative to the level we observe in 2017, search activities were close to zero during the first few years of ride-hailing entry (2010-2012). In particular, search activity for Uber only begins to rise after UberX is officially announced in 2013. We plot the same indices for Austin, Texas in **Figure 2-(b)**, which provides a strong case for our argument. In May 2016, both companies officially announced their exit from Austin after the city's voters supported fingerprint checks for their drivers. But, they returned a year later after the state passed the bill that requires the minimum background checks but no fingerprint checks for their drivers (L.A. Times, May 29, 2017). The figure confirms that both Uber and Lyft search activities rose sharply at the same time after their entry, but collapsed immediately after both companies announced their exit upon the fingerprint ruling in May 2016, and returned to the pre-2016 level after the passing of the state bill that overturned the fingerprint rule. Furthermore, although we do not report on other cities, we observe all sorts of patterns in these indices: Search activities exist for both Uber and Lyft in some cities, only one in some, and none in others; Search activity for Uber precedes Lyft in some cities and vice versa in other cities. For these reasons, we believe Google Trends indices for Uber and Lyft in each city are good proxies for their market presence in that city.

Given the above, we construct a unified measure of ride-hailing entry based on these normalized Google Trends indices as follows. For each MSA in our sample, we obtain search trend indices from January 2010 to December 2018, using "Uber" and "Lyft" as keyword entries.[7] We then define entry if the maximum of the two indices exceeds a certain threshold. That is, for each MSA i and for each month t,

$$Entry_{it} = \boldsymbol{I}\left\{\max(Trends_{it}^{Uber}, Trends_{it}^{Lyft}) \geq c\right\}.$$

For all subsequent analyses, we use the cutoff value of $c = 0.2$. We tested several values for the cutoff, but eventually chose the cutoff value of 0.2 to match the ideal of our quasi-experimental design. That is, the cutoff value must be such that below the cutoff, there is virtually no Uber/Lyft activity and above the cutoff, Uber/Lyft activity jumps and continues to grow thereafter. Put differently, we avoid the cutoff values that would generate similar Uber/Lyft activity levels in both sides of the cutoff.

Another problem is that even in cities where no ride-hailing exists, residents may still search for Uber or Lyft for other purposes. Because the indices are normalized against the highest month of search trends for each MSA, even low search frequencies can result in a large search index. This may result in the false signal for ride-hailing entry. We correct for

[7] Google Trends are calculated for digital marketing areas (DMAs). Hence, we use 2009 MSA boundaries to convert DMA boundaries to MSA boundaries.

such anomaly in the following way. First, for each MSA, we calculate the maximum search index observed *before* December 2012 — i.e., before any entry would have occurred for most cities. We then subtract this pre-treatment maximum from the original index values. This normalization ensures that the adjusted index values exceed zeros only if their original values exceed the pre-treatment maximum. Second, we hire a graduate student to manually search for fictitious rides for Uber and Lyft on randomly chosen locations in each city using the smartphone's google map application. If we don't find any UberX or Lyft service for the fictitious ride requests, we classify that city as having 'no service' as of 2020. We then visually check each city individually and confirm that for no-entry MSAs, the normalized trend indices never cross the threshold whereas for entry MSAs, the indices continue to grow over time and stays above the threshold once they cross it.

Panel (a) of Figure 3 compares the distribution of our *de facto* entry dates constructed from this measure against that of (i) *reported* entry dates of UberX, taken from Hall *et al.* (2018) and (ii) *official* UberX launch dates obtained from Uber Inc.[8] We see that both reported entry and official entry precede *de facto* entry for some cities while a much larger number of MSAs are recorded with ride-hailing entry in our *de facto* entry measure. The differences also come from the fact that we account for Lyft's activity. Thus, readers should be cautious in comparing our study with earlier studies because we use a different measure of ride-hailing entry (and a different identification strategy as a consequence).

Panel (b) of Figure 3 compares the event-study estimates of the effects of entry on the normalized Google Trend Index, using the three entry measures.[9] There are several important take-away messages from the figure. First, the estimated impacts of reported entry and official entry are quite small over the two-year window after entry, and even smaller than the cutoff value ($c = 0.2$) used to determine the *de facto* entry. Thus, if we use reported entry as the treatment variable, we would only get at the average treatment effect from this small activity level of Uber/Lyft. Second, there is a discontinuous jump in the estimated impact at the *de facto* entry timing, suggesting that the Google Trend Index increases rather abruptly around the cutoff value. We take this as a sign that our choice of the cutoff value is well calibrated to mimic the quasi-experimental design. Third, the estimated impacts of *de fact* entry gradually increase over the two-year window. This implies that we would expect the treatment effect of Uber/Lyft entry to be dynamic, gradually increasing over time. Fourth, there is no sign of violation of the no-anticipation assumption even if we use the *de facto* entry as our treatment variable — the estimated impacts during the

[8] Our special thanks go to Michiko Numazu, a data scientist at Uber Inc. who kindly provided the data for us.

[9] We use the dCDH/CS estimator without covariates.

pre-entry periods are nearly zero and are precise.

B. Ambient Air Quality: Our second data source is the satellite-based nitrogen oxides (NO_2) data from the National Aeronautics and Space Administration (NASA)'s Goddard Earth Sciences (GES) Data and Information Services Center (DISC). GES-DISC provides level-3 daily total column NO_2 in molecules per cm^2 on the 0.25×0.25 degree global grids.

We focus on ambient NO_2 pollution not only because it is an important transport-related air pollutant but also because we wish to minimize the risk of falsely capturing the effects of other confounders on air pollution. NO_2 is a leading cause of respiratory diseases such as asthma, is a known precursor to ozone (O_3), and is often used as the indicator for the larger group of nitrogen oxides (NO_x). Because NO_2 is released into the air from combustion of fossil fuels, the primary sources of NO_2 emissions are cars, trucks and buses, power plants, and off-road industrial equipment. According to the U.S. Environmental Protection Agency (EPA) National Emissions Inventory (NEI), road transportation accounts for roughly 36% of total NO_2 emissions, but accounts for only 2-4% of $PM_{2.5}$, PM_{10}, and VOC emissions in the United States. The share of road transportation gets even higher in urbanized areas. For example, road transportation accounts for 48% in Chicago. Therefore, if there is any effect of ride-hailing on transport-related air pollution, we should expect to see it most vividly in NO_2 concentration levels.[10]

We rely on the satellite-based air quality data for its coverage, granularity as well as temporal consistency. AirData from U.S. EPA provide daily air pollution data from monitoring stations at the ground level for all five criteria pollutants as well as climate data such as temperatures and wind speeds for all monitoring stations in the U.S. However, the monitoring stations are only sparsely located across U.S. cities. The monitoring data are also sparsely monitored over time, with missing records for some months and with monitoring stations frequently added and discontinued. As a result, ambient air pollution data are consistently available over time only for 69 MSAs (18.8% of the full MSA sample). Our main analysis would thus have to be restricted to this subset of MSAs if we rely on monitoring data. Furthermore, recent studies find that air quality data from monitoring sites may be systematically biased due to strategic compliance behavior, either by local authority (Grainger *et al.*, 2021) or by polluters (Zou, 2021).

For illustration, **Figure 4** plots our satellite-based NO_2 grid data along with monitoring site locations for three example cities in the U.S. The NO_2 data are grid-level monthly averages for January, 2014. The map also shows the MSA boundary as well as the Census-

[10]Previous studies have also used carbon monoxide (CO) and aerosol optical depth (AOD) for similar reasons. But we chose not to do so because we only have access to the satellite-based CO and AOD data on much less granular level (0.5×0.5 grided cells or larger).

tract-level population density. The monitoring sites are located mostly around the city centers, but only sparsely. As a result, the monitoring sites do not necessarily coincide with high-pollution areas. Furthermore, the map shows that the high-pollution incidence can go beyond the central urban areas to suburban or non-urban areas. This occurs partly due to wind direction but also due to suburb-to-urban or suburb-to-suburb commuting. Our argument outlined in **Section 2** suggests that ride-hailing entry may have different air-quality impacts for different subareas of cities, say, urban versus non-urban areas. Another benefit of using the satellite-based data is that it allows us to calculate (area-weighted) monthly means and maximums for urban, suburban, and non-urban areas within each MSA. We use the 2010 Census definition of urban-area boundaries for all study periods. To validate the satellite-based data, we regress these satellite-based means and maximums on EPA's monitoring data. Consistent with earlier studies, the satellite-based data are highly significantly correlated with the monitoring data, with the urban-area statistics having the most predictive power. See also **Figure A3** in the **Appendix,** which plots the satellite-based NO_2 data against the EPA monitoring data using MSA-level monthly averages.

C. Commuting Patterns and Other Variables: We also use yearly household-level commuting mode data in order to explore the economic mechanism behind our main results on NO_2 concentration levels. To do so, we compile all microfiles from the U.S. American Community Survey (ACS) to construct repeated cross-section data. We only have panel structures at the MSA level since the microfiles do not allow us to trace out household identifiers over years to construct panels at the household level. We also use the ACS data to construct a variety of pre-treatment covariates such as the median age, the median income, the share of manufacturing employment, and the share of college graduates.

We also supplement our data from a few other sources. We obtain monthly temperature and wind data from the EPA's Air Data, monthly regional gasoline price data from North American Electric Reliability Corporation (NERC), MSA-level population density from U.S. Census Bureau, and historical time series of county-level nonattainment status for five criteria pollutants from EPA's Greenbook website. We use these to construct time-varying control variables.

As discussed in **Section 3**, the parallel-trend assumption may be violated if there exist confounding factors that would systematically affect the trends in NO_2 concentrations that are also correlated with the location and timing of Uber/Lyft entry. One plausible confounder is the local compliance efforts for the National Ambient Air Quality Standards (NAAQS). The 1990 Clean Air Act Amendments requires the EPA to set the NAAQS for five criteria pollutants. Under the CAAA, if a county is designated as "nonattainment" of the standards, then states were mandated to regulate plant-level sources of these pollutants. Although the

NAAQS stayed the same for NO_2, major changes to the NAAQS occurred in 1997, 2006, and 2012 for PM2.5 and in 1997, 2008, and 2015 for O_3. Because NO_2 is a known precursor to both pollutants, the nonattainment status for either pollutant may affect the time trends for NO_2 concentrations at the MSA level. Thus, we control for these in our regression.[11]

D. Descriptive Statistics and Treatment-Control Structures: **Table 1** provides a statistical overview of main outcome and pre-treatment variables by treatment status. The earliest *de facto* entry occurs in 2013 whereas the majority of entry occurs in 2014. Roughly 18% of the MSAs in our sample do not show sufficient Uber/Lyft activity to be classified as *de facto* entry. The table confirms that entry is correlated with MSA-level socioeconomic covariates: Uber/Lyft enter cities with higher population density, higher median income, higher share of college graduates, and higher share of public transit commuters earlier than others. Furthermore, there is an indication that cities with early Uber/Lyft entry may be more polluted, as measured in EPA monitoring data. However, the use of satellite-based NO_2 data reveals that this may be due to systematic bias in monitoring data. EPA monitoring data are available for about 22-27% of observations for cities where Uber/Lyft enter before 2015 whereas only 5-8% are available for cities where Uber/Lyft entry never occurs or occurs after 2015. If we use the satellite-based data, differences in NO_2 concentrations across cities by entry status become less obvious.

5. Estimation Results

5.A. TWFE Estimates

We start by presenting the estimation results from the TWFE regression of eq. (1).[12] We use Sergio Correia's *Reghdfe* package in Stata to efficiently absorb multi-way fixed effects. **Table 2-A** presents the results using (the area-weighted means of) monthly means (in logged values) as the outcome whereas **Table 2-B** uses (the area-weighted means of) monthly

[11] There is a subtle discussion as to whether the nonattainment statuses for these pollutants are 'bad controls' in our context. On one hand, the local authorities with nonattainment status for PM2.5 or O_3 may take a variety of compliance measures, some of which may affect the local emissions sources of NO_2 directly or indirectly. Hence, the *failure* to control for them is likely to *overstate* the estimated impact of ride-hailing entry if the nonattainment status coincides with the entry timing/location. On the other hand, ride-hailing's impact on ambient NO_2 may also affect ambient levels of PM2.5 or O_3, which in turn may trigger changes in the nonattainment statuses for these pollutants. If this were the case, the *inclusion* of these controls may lead to the *overstatement* of the entry's impact. As we discussed above, transport-related emissions account for only 2-4% of PM and VOC pollutants. Hence, the latter effect must be quite small even if it exists. Our estimation results seem to confirm this point — the estimated impacts of entry on ambient NO_2 are smaller when we include these controls.

[12] We use Sergio Correia's reghdfe package in Stata to efficiently absorb two-way fixed effects.

maximums (in logged values). Each table shows the results from three regressions, with varying sets of controls, for each of the subareas (urban, suburban, non-urban). The first column controls for MSA/year/month fixed effects, the second controls for climate conditions, and the third includes other time-varying covariates such as MSA-specific linear time trend, regional gasoline price, and nonattainment status for O_3 and $PM_{2.5}$. Inclusion of endogenous time-varying controls can potentially bias our estimates since they can be correlated with unobservables in the estimating equation. Hence, we avoid use of such controls. Standard errors are clustered at the MSA level.

In both tables, there is a tendency that the estimated impacts tend towards positive values with more controls, suggesting that our estimates may be biased towards negative values. Hence, we focus on the results with full controls. In **Table 2-A**, we see that *de facto* ride-hailing entry is estimated to decrease ambient concentration levels of NO_2 for cities in the highest population density quartile. The estimates are statistically significant, and range from -0.032 log points (in non-urban areas) to -0.034 log points (in urban areas). If we use a linear transformation evaluated at the mean concentration levels, these estimates imply that ride-hailing entry reduces NO_2 concentrations by 0.50-0.53 ppb for these cities. Interestingly, the estimates get smaller in magnitude for less densely populated cities, and eventually turn positive for cities in the lowest density quartile. The results are consistent with our prediction in **Section 2** — In cities with sufficient public transit networks, a combined use of ride-hailing with public transit can displace use of private cars, reducing overall vehicle emissions, but in cities without such public transit networks, ride-hailing may simply induce more driving, increasing overall vehicle emissions.

Next, we turn to the results on monthly maximums in **Table 2-B**. The results are generally consistent with those of **Table 2-A.** That is, the estimates are negative and statistically significant for cities in the highest density quartiles; the estimates get smaller for cities in the lower density cohorts, and turn positive for non-urban areas in the lowest density cities. There is one important difference, however. The estimated impacts are larger, more statistically significant, and range from -0.049 log points (in suburban areas) to -0.050 log points (in urban areas). From these, we infer that ride-hailing entry mostly affects the peaks of transport-related air pollution. That is, it tends to reduce transport-related air pollution on days when pollution levels are high (i.e., when heavy vehicle traffic are expected). This also explains the relatively large estimates on the mean concentration levels because the monthly means are highly sensitive to the maximum concentration levels. These results are also consistent with Kim and Sarmiento (2021), who finds that much of the air qualitydCDH improvement comes from a decline in the number of bad air quality days in summer.

5.B. Event Study Estimates

Next, we present the event study estimates from the dCDH/CS estimator.[13] As explained in **Section 3**, we implement the dCDH estimator separately for each density quartile (rather than pooling all observations) after purging out the influence of exogenous time-varying covariates. Hence, the identification is much clearer here. By construction, this estimator uses only 'not-yet-treated' MSAs for each treatment timing within the same population density cohort as comparison units. As a result, the estimates are not contaminated from use of either already treated MSAs or other non-similar MSAs in different density cohorts.

Figure 5 plots the estimates of the ATTs in relative time to entry ($\tau = 0$). The first two graphs on the top panel plot the estimates pooling all MSAs. The other graphs plot the estimates using a subsample of MSAs on each population density quartile. To avoid busy graphs, we only present the results using monthly maximum NO_2 except for the first graph on the pooled sample. For the highest density cohort, eventually all units get treated, and hence, we are able to estimate the ATTs only within the two-year window. Thus, for consistency, we use only not-yet-treated units as control units and estimate the ATTs over the two-year window for all density cohorts.

We see that in line with the results from the TWFE regression, on the pooled sample as well as on the highest density cohort, NO_2 concentrations start to decline after ride-hailing entry (relative to not-yet-treated units), and the magnitudes of the decline get larger over time. This is also consistent with **Figure 2**, which shows that Uber/Lyft activity grows over time. On the other hand, we do not see any sign of either a rise or decline in NO_2 concentrations (relative to not-yet-treated units) for the medium- and low-density cohorts. The graphs also indicate that there is no sign of violation of parallel-trend assumption during the pre-treatment period. There results boost our confidence in the estimated impacts of ride-hailing entry.

6. Exploring Economic Mechanism

So far, our results are consistent with the economic mechanism outlined in **Section 2**. That is, while ride-hailing may increase congestion and vehicle emissions by inducing deadheading or displacing of mass transit for parts of daily trips, it may still decrease overall

[13] We use *multiplegt* in Stata to implement the dCDH estimator. The CS estimator (using *csdid* package in Stata) produces essentially the same results, but was far slower in our computing environment.

air pollution if a combined use of ride-hailing with mass transit displaces private car use more than such adverse behavior. Thus, transport-related air pollution may decline in cities with high public transit density where the complementarity between ride-hailing and public transit is strong. In this subsection, we explore whether such an economic mechanism indeed exists behind our results.

To do so, we draw household-level data from the American Community Survey (ACS) on workers' commuting mode choice. We construct an indicator I_{ijt} of each household i's commuting mode j to work in year t, and run the TWFE regressions similar to eq. (1), with I_{ijt} as outcome variables. We run the regression separately for each commuting mode j. We do not impose any structural restriction on the parameters across these regressions. Because we only have access to PUMA data (so our geographic identifiers are county and city of residence), we only have panel structures at the MSA level. Hence, we include household-level demographic controls such as education, income, and race.

In the ACS survey, respondents are asked to record only one commuting mode they used to get to work during the week before the survey, and if they use more than one method, they are asked to record one that is used for most of the distance. The ACS's choice set includes 12 methods of transportation. For ease of interpretation, we consolidate these into three primary commuting modes for the first set of regressions: commuting by private car (I_c), commuting by public transit (I_{pt}), and commuting by other modes (I_o). These three modes account for 99.1% of all commuting mode choices in our sample (85.6%, 4.8%, 8.7%, respectively for I_c, I_{pt}, and I_o). The remaining 1% of the sample commutes by taxicab, motorcycle, or bicycle.[14] In the second set of regressions, we disaggregate 'other modes' further into three modes: commuting by walking (I_{walk}), working at home (I_{home}), and other (I_{other}). Thus, $I_o = I_{walk} + I_{home} + I_{other}$ by definition.

The difficulty we have is that none of the transport modes in the ACS survey directly captures the use of ride-hailing services. Our argument here is that $I_{pt} + I_o$ is a good proxy for the use of ride-sharing, $I_{pt} + I_o \approx I_{rs}$, for several reasons. The starting point for our argument is that the ACS survey format does not allow respondents to record the combined use of several transport methods, and hence, some of the actual daily transport choices are likely to be absorbed into one of the chosen transport modes. On one hand, commuters using ride-hailing mostly to connect to public transit stations are likely to appear as an increase in I_{pt} whereas commuters using ride-hailing mostly for daily errands would show up as an increase in I_{walk}. On the other hand, only those who use ride-hailing as a primary

[14] There is a possibility that Uber/Lyft may be synonymized with conventional taxi, and the commuters using Uber/Lyft to work may record 'taxi' instead of 'other' as their primary commuting mode. Hence, we also estimated the same regressions in **Table 3 and 7** using $I_o = I_{walk} + I_{home} + I_{taxi} + I_{other}$ instead, but the results were quite similar.

commuting mode (e.g., car-pooling or using ride-hailing directly to the employment center) would appear in I_{other}. This line of reasoning suggests that much of the increase in the use of ride-hailing might be *unobserved* by nature. However, there must be an increase in the supply of ride-hailing to accommodate such an unobserved increase in the use of ride-hailing (i.e., associated with increases in I_{pt}, I_{walk}, or I_{other}). We would expect this would show up as an increase in I_{home}. Admittedly, this is an important limitation to the use of the ACS data for our study purpose, and hence, we refrain from making strong inferences from the ACS data.

With this limitation in mind, we estimate the impact of ride-hailing entry on commuting mode indicators I_{ijt} and see if the estimates are consistent with the economic mechanism we discussed in **Section 2**. Per our discussion, we expect vehicle emissions to increase if commuting by public transit declines ($I_{pt} \downarrow$) and commuting by ride-hailing increases ($I_o \uparrow$) while commuting by private car stays roughly constant ($\Delta I_c \approx 0$). This is the case of ride-hailing being a substitute for public transit. In contrast, we expect vehicle emissions to decrease if both commuting by public transit and by ride-hailing increase ($I_{pt} \uparrow$ and $I_o \uparrow$) while commuting by private car declines ($I_c \downarrow$). This is the case of ride-hailing and public transit being complements to each other, and a joint substitute for commuting by private car. An ambiguous case occurs when commuting by private car declines ($I_c \downarrow$) and commuting by ride-hailing increases ($I_o \uparrow$) while commuting by public transit declines ($I_{pt} \downarrow$). There are, of course, other plausible explanations, given the limitation of our measurement on I_{ijt}. We discuss this point below when discussing the results.

Table 3-A reports the results of the TWFE regressions on the first set of outcomes (I_c, I_{pt}, and I_o). In all regressions, we include the MSA and year fixed effects as well as household-level demographic controls. For each outcome, the second column includes weather controls whereas the third column also includes other time-varying controls such as gasoline price and nonattainment status. **Table 3-B** repeats the same, but use the sub-categories of I_o as outcomes (I_{walk}, I_{home}, and I_{other}). As with our main results, there is an indication that the estimates get smaller in magnitude as we include more controls. Hence, to be conservative, we focus on results with full controls.

The results in **Table 3-A** are not only consistent with our main results in **Table 2** but also with the economic mechanism discussed above. In the highest density MSAs, the ride-hailing entry is estimated to decrease the share of commuting by private car by 1.3 ppt, increase that of public transit by 0.4 ppt, and increase that of other modes by 0.7 ppt. The estimated impacts roughly sum to zero, and hence, the estimated impacts are consistent with each other (despite that we make no structural restriction on them). As discussed above, the estimates for the highest density MSAs correspond to the case of ride-hailing and public transit being complements to each other while being a substitute for private car commuting. Hence, these estimates are also consistent with the pollution-decreasing effect of ride-hailing entry for these MSAs in **Table 2**. Furthermore, in the third-quartile MSAs, the estimates are significantly negative for private car commuting, insignificant for public transit commuting, and significantly positive for other modes. This corresponds to the case of ambiguous pollution impacts. For the second quartile MSAs, we essentially see no impact of ride-hailing on commuting modes. For the lowest density quartile, we see that ride-hailing entry is associated with increases in private car commuting. All of these are indeed consistent with the signs of the effect of ride-hailing entry reported in **Table 2**.

Although the results so far are consistent with our expectation, turning to the results in **Table 3-B** reveals something we did not expect a priori. In the highest density MSAs, ride-hailing entry is significantly associated with increases in commuting by walking (I_{walk}) and working at home (I_{home}), but has no effect on the other mode (I_{other}). As discussed above, this may imply that much of the use of ride-hailing is absorbed into the combined use of ride-hailing with either public transit or walking. We also interpret the increase in the share of those working at home as the sign of the associated increase in ride-service providers. We, however, found it somewhat puzzling to see there is no impact on the other commuting mode. A priori, we expected I_{other} to be a good 'direct' indicator of the use of ride-sharing service, and hence, we expected its share to increase. But, the result suggests that this was not the case, and instead, the ride-hailing entry did not change the share of those who use ride-sharing or carpooling as the primary commuting mode. Taken together, these results seem to suggest that commuters are using ride-hailing service mostly as a complementary means of commuting rather than the primary commuting mode.

Interestingly, in the third quartile, virtually all of the increase in the share of other modes is explained by the increase in the share of those working at home, and the magnitude of the impact (+0.5 ppt) is roughly the same as that of the reduction in the share of commuting by private car (-0.4 ppt). On one hand, these two effects tend to cancel out each other in terms of their impact on *mean* NO_2 concentrations in these MSAs because the shares of the other commuting modes did not change. On the other hand, these effects tend to decrease *maximum* NO_2 concentrations in these MSAs because private car commuters decline while commuters using ride-hailing service as the primary mode do not change — ride-hailing service providers are mostly offering services to those who use the services as complements to their primary modes of travel. Hence, these results are consistent with findings from Hall *et al.* (2018) and Kim and Sarmiento (2021).

7. Alternative Identification Strategies

The results so far rely on the conditional parallel trend assumption **A1** to hold on each population density cohort. Because we include a MSA-specific linear trend, this identifying assumption would be valid as long as there are no unobservable shocks that happen to occur at the same timing as the ride-hailing entry that are correlated with the outcomes in the same cohort of MSAs. Here, one disadvantage of our empirical strategy arises from the fact that we construct our *de facto* entry measure from the Google Trends Indices. The Google Trends Indices are the equilibrium outcomes, rather than the measures of exogenous changes in supply of ride-hailing service (Hall *et al.*, 2018). Suppose, for example, highly dense cities consist of a larger share of residents who have higher demand for more eco-friendly means of transportation. The changes in their behavior over time may be correlated with changes in demand for ride-hailing services. In such a case, the estimated impacts in the previous section get at the combined effect of their demand for eco-friendly transportation and that for ride-hailing. To address this issue, we take two alternative approaches. The first is to exploit quasi-experimental settings in Texas, and the other is to apply the instrumental variable (IV) approach using geography-based instruments.

7.A. Quasi-experimental Setup in Texas

In this approach, we restrict our study sample only to twenty-four MSAs in Texas. There are several advantages for doing so. First and foremost, Austin, the state capital of Texas, voted for the fingerprint rule (FR) in 2016, which would mandate the fingerprint check for all ride-hailing drivers. In response to the rule, Uber and Lyft halted all their services in May, 2016. A year later, both companies resumed the services at the end of May, 2017 after the state passed the bill that removes the requirement for fingerprint checks for their drivers. As indicated in **Figure 2-(b)**, this incident caused a sharp, unanticipated decline and a subsequent rise in the supply of Uber/Lyft services. Importantly, this incidence occurred only in Austin, but no other neighboring MSAs. There is one caution here — a number of small ride-hailing companies, such as GetMe and RideAustin, operate in Austin during this period. Therefore, Uber and Lyft customers may have switched to their service. Hence, we take the FR as an event that caused an exogenous change in the supply of ride-hailing service, but not the complete exit. Second, *de facto* entry dates mostly correspond to *reported* entry dates in Texas. Therefore, the treatment assignment is unlikely to be driven by unobservable demand shocks that are correlated with the outcomes. This minimizes the risk of identification failures we discussed above. Third, there is sufficient variation in treatment status. Out of twenty-four MSAs, entry occurs in fourteen MSAs in 2014 and in two MSAs in 2016. We do not observe ride-hailing activity for the remaining eight MSAs. The urban areas of these MSAs are sufficiently far apart from each other (see **Figure A4**). Therefore, focusing on Texas allows us to control for state-level confounds while exploiting variation in timing and location of ride-hailing entry.

There are also disadvantages, however. First, the small sample size precludes the use of certain estimation strategies for reliable statistical inference. In particular, we cannot reliably use the dCDH/CS estimator because there are not enough 'relevant' MSAs for each treatment timing for each density cohort. Second, air pollution concentration data are highly volatile, with large variation remaining even after controlling for many observables. This also prevents us from using certain (otherwise credible) approaches such as regression discontinuity (RD). For example, Tarduno (2021) also uses Austin's FR as a natural experiment and estimates the difference-in-differences RD regression, comparing hourly traffic speeds before and after May 9th, 2016 against those of 2015. Unfortunately, as we shall see below, this approach fails to deliver credible statistical inference in our case.

Given these pros and cons, we employ two estimation strategies. Our first strategy is to apply a version of the synthetic control method (SCM) à la Abadie *et al.* (2010) to visually present the impacts of both the ride-hailing entry and the FR relative to the counterfactual trend. We implement this strategy as follows. We first estimate the auxiliary regression (2) using only the pre-treatment observations and predict the residuals for all observations on the Texas sample. We then apply the SCM using lags up to 12 months as covariates. By construction, this allows us to use all available information to construct the counterfactual Austin's residualized NO_2 trend. **Figure 6** plots the synthetic Austin's NO_2 residuals (squares in maroon) against the actual Austin's NO_2 residuals (dots in navy). To visualize the impacts, we also plot the local polynomial smoothing of these residuals, respectively as dashed and solid lines. **Figure 6-A** uses monthly means whereas **Figure 6-B** uses monthly maximums. The figure indicates that both mean and maximum NO_2 concentrations have an upward trend before the ride-hailing entry *even after controlling for observable confounds*, but switch to a declining trend after the entry. The declining trend continues during the FR period, and then returns to the pre-FR level after the FR period. Importantly, we see these trends relative to the counterfactual NO_2 trends, not just relative to the pre-treatment period. The figure also illustrates why the RD design is unlikely to work in our case — we see unusually high (non-seasonal) NO_2 concentration levels to the right of both the ride-hailing entry and the FR cutoffs, even after controlling for year/month fixed effects.

Our second approach is to estimate the TWFE regression in (1) on Texas MSAs while adding another FR dummy, which equals 1 for Austin during the FR period. This approach does not take full advantage of the sharp identification due to the FR, but allows us to explore its impact on the commuting mode choices. **Table 4** presents the results on NO_2 concentrations. To make the results comparable to **Table 2**, we use the same population density quartiles as in **Table 2**. The results are somewhat similar to those of the full sample reported in **Table 2**, with notable exceptions. The estimated impacts of ride-hailing entry are negative and statistically significant for MSAs with higher population density, both on monthly means and monthly maximums. As in **Table 2**, the estimated impacts are generally larger on monthly maximums than on monthly means. Unlike in **Table 2**, however, the estimated impacts on the second quartile MSAs are positive and statistically significant when monthly maximums are used. Interestingly, the fingerprint rule is estimated to decrease NO_2 concentrations, and the estimated impacts are large. Note that the FR dummy is defined as the Austin dummy times the FR period dummy. Hence, the impacts are estimated relative to no-entry status. Because Austin belongs to the highest population density quartile, we obtain the pure impacts of the FR by subtracting the estimates for that quartile: 0.051-0.056 log point decrease in mean NO_2 concentrations. This result may seem puzzling at first because it implies that both the increase (i.e., ride-hailing entry) and the decrease (i.e., FR) in the supply of ride-hailing service reduce transport-related air pollution. However, exploring the economic mechanism behind the results, we find that they are indeed consistent, not only with our argument, but also with each other.

Table 5 reports the results of our supplementary TWFE regressions using yearly household-level commuting modes as outcomes. The same explanations as in **Section 6** apply to all regressions in **Table 5**, except that the regressions are run on the Texas sample along with the additional FR dummy. The results suggest that the FR is estimated to decrease private car commuting by 1.1 ppt, increase public transit commuting by 0.5 ppt, and increase other commuting modes by 0.6 ppt. These changes are also associated with a 1.2-ppt increase in the share of those working at home, a 0.4-ppt decrease in commuting by walking, and a 0.2-ppt decrease in the other mode. By subtracting the estimates for the highest population density quintile, we arrive at the pure impacts of the FR in Austin (relative to the pre-FR period): -0.5 ppt in private car commute, +0.6 ppt in public transit commute, -0.6 ppt in walking to work, +0.5 ppt in working at home, and -0.4 ppt in the other mode. Thus, we conclude that the sharp decline in the supply of ride-hailing service due to the FR further decreased commuting by private car while encouraging the use of public transit, and that these changes in the commuting patterns decreased transport-related air pollution. Although the exact economic mechanism is still ambiguous, the results are consistent with our argument in that the impact of ride-hailing depends on how it changes the overall commuting patterns.

7.B. Instrumental Variable Approach

Our last identification strategy is to apply an instrumental variable (IV) approach to the DD regression (1), building upon recent advances in empirical research on economic geography. To arrive at credible instruments, we distinguish two sources of endogeneity. Our entry variable $D_{c\tau}$ is an indicator of *when* and *where* our normalized Google Trends Index crosses the threshold. Because the frequency of the keyword searches is an equilibrium outcome, both the *level* and the *timing* of treatment is endogenous and may coincide with unobservables that affect air quality. Therefore, we need instruments that predict both the *location* and the *timing* of ride-hailing service's market penetration, yet do not affect ambient air pollution directly after controlling for pre-treatment covariates.

To arrive at our instruments, we first start by constructing three types of geography-based instruments [see Redding and Turner (2015) for a more thorough discussion of these instruments]. The first is based on the highway construction plan as of 1947 developed under the mandate to serve military services. This IV is known as the "planned route IV" and is used in Baum-Snow (2007) and Duranton and Turner (2011; 2012). The second is based on the railroad network as of 1870. This IV is known as the "historical route IV" and is used in Duranton and Turner (2011; 2012). The third is based on the Euclidean spanning network connecting large cities as of 1860. This is known as the "inconsequential unit IV" and is used in Faber (2014). All three variables are the MSA-level indicators, each of which equals 1 if the route passes through the MSA.

Although these IVs exploit conceptually quite different quasi-random variations, they are often used in conjunction with each other and generate quantitatively similar variations in practice (Redding and Turner, 2015). In essence, these IVs are designed to construct 'hypothetical' routes for a given geography, which predict the observed routes, but are presumably uncorrelated with contemporaneous economic shocks after controlling for the current economic conditions. These geography-based instruments are known to work well for estimating the causal effect of road network on economic outcomes. We tailor this line of argument for our purpose: the 'hypothetical routes' are a good predictor of the current population size, which is a good predictor of ride-haling entry/activity, but do not affect the current air quality except through the ride-hailing entry, after controlling for the current (pre-treatment) economic conditions. Hence, these variables serve as a set of good instruments for the *location* of ride-hailing service's market penetration. Details of how we construct these instruments are explained in the **Appendix F**.

The geography-based instruments, however, give us only cross-sectional variation, and thus, are a poor predictor of the *timing* of the market penetration. For this, we use *reported* entry dates of UberX and Lyft from Hall *et al.* (2018), which are collected from newspapers, blog posts, and web entries, and hence, are plausibly considered official entry dates of these services. Specifically, we construct our IVs as follows. For each geography-based instrument k, $Z_{c\tau}^k$ for city c and time τ is:

$$Z_{c\tau}^k = R_{c\tau} \times G_c^k.$$

On one hand, because ride-hailing service's market penetration into a city occurs only after the service officially enters the city, *de facto* entry dates are monotonic in reported ride-hailing entry dates ($D_{c\tau} \geq R_{c\tau}$). On the other hand, as Hall *et al.* or Berger *et al.* (2018) argue, the timing of reported ride-hailing entry is exogenous after controlling for pre-treatment covariates such as the current population size. Identification is strong because we interact reported entry with geography-based instruments — we only use the timing of reported entry to predict the timing of *de facto* entry whereas using hypothetical routes to predict the location of *de facto* entry.

For implementation, we apply the two-stage least squares (2SLS) estimator on the DD regression in eq. (1), separately on each population density cohort (i.e., not pooling all MSAs). We do this to ensure we can interpret our estimates as the local average treatment effect (LATE). For the DD-IV estimates to be the valid LATE, we require the following identifying assumptions (Duflo, 2001; Hudson *et al.*, 2017):

A2. *For MSAs in the same population density cohort* s*,*

1. **Relevance and monotonicity**: $E[D_{c\tau}|Z_{c\tau} = 1] \geq E[D_{c\tau}|Z_{c\tau} = 0]$.
2. **Exclusion restriction**: $Z_{c\tau}$ *affects* $Y_{c\tau}$ *only through* $D_{c\tau}$.
3. **Conditional parallel trends in $Y_{c\tau}$ on $Z_{c\tau}$**: *Conditional on exogenous (time-varying) covariates* $\Omega_{c\tau}$*, parallel trends in unobservables hold for all MSAs in the same population density cohort* s*.*

$$E[\epsilon_{c,\tau} - \epsilon_{c,\tau-1}|Z_{c\tau} = 1, \Omega_{c\tau}] = E[\epsilon_{c,\tau} - \epsilon_{c,\tau-1}|Z_{c\tau} = 0, \Omega_{c\tau}], \quad \forall\ c \in s,\ \forall\ s \in \mathcal{S}.$$

Table 6 reports the DD-IV estimation results on mean NO_2 concentrations, using all three geography-based instruments interacted with the reported ride-hailing entry dummy. We use the same set of controls and the standard errors are clustered at the MSA level as in **Table 2**. Our DD-IV estimates are robust to varying sets of controls. Hence, we only report the estimates with full controls. In the table, we also report Cragg-Donald's F statistics for weak IV as well as Hansen's J statistics for over-identification (for urban-area regressions only). The F statistics are well above Stock-Yogo's critical values, and J statistics are not statistically significant. There is a sign, however, that the DD-IV estimates are less precise than the TWFE estimates. This is expected not only because the 2SLS estimation generally leads to larger standard errors but also because we run the 2SLS estimation on each density subsample.

The table indicates that the DD-IV estimates have essentially the same signs as the TWFE estimates, but are much larger in magnitude. Ride-hailing entry is estimated to reduce mean NO_2 concentrations for cities with high population density (i.e., the 4th quartile), particularly in urban areas. The signs of the estimates turn positive for medium- and lowest-density MSAs. The estimated effects on the 2nd quartile MSAs are negative but statistically insignificant due to their large standard errors (consistent with the TWFE regression). We take these as supportive evidence for our main results.

We also run the DD-IV regressions on household-level commuting modes, in the same way as in **Table 6**. The results are reported in **Table 7**. The DD-IV estimates are quite similar, at least in signs, with the TWFE estimates in **Table 3**, and are also consistent with the estimated impacts on air quality in **Table 6**. For the highest density MSAs, ride-hailing entry is estimated to decrease private car commuters by 6.1 ppt, but much of this decline is associated with an increase in public transit commuters (by 5.0 ppt) rather than in other modes. In contrast, for the 3rd quartile MSAs, a decrease in private car commuters (by 1.3 ppt) is accompanied by an increase in the share of other commuting modes (by 0.9 ppt), and much of this change comes from an increase in the share of working at home (by 0.6 ppt).

The magnitudes of the estimates seem economically too large, however, particularly for lower population density cohorts. For example, ride-hailing entry is estimated to increase mean urban-area NO_2 concentrations by 0.121 log points in the lowest quartile and to decrease by 0.119 log points in the second quartile. There can be two interpretations on this result. The first is, of course, that our instruments satisfy all identifying assumptions in **A2**, and thus, our DD-IV estimates get at the true LATE estimates for the 'compliers': i.e., the true impacts on MSAs that are affected by the instruments. The second interpretation, however, is that our instruments fail to satisfy one or more assumptions in **A2**, and thus, our estimates are biased. We are lean toward the second interpretation for the following reason. Recall that our treatment variable is a binary indicator of entry, and hence, our DD-IV estimand can be written in a manner similar to the familiar Wald formula. Hence, if our instruments do not predict entry well, then it tends to overly inflate the estimated impacts. Now, recall our discussion on reported entry (or see **Figure 3-(a)**). Reported entry dates are incomplete, and their coverage is particularly poor for small, low-density cities. This is evident in relatively small Cragg-Donald's F statistics in **Table 6**. Therefore, our DD-IV estimates are likely biased away from zero, particularly for lower-density MSAs, although the signs of the estimates should stay the same as long as the other identifying assumptions hold. Given this, we refrain from drawing too strong conclusions, particularly on lower density cohorts. Nonetheless, ride-hailing entry seems to affect commuting patterns quite differently across different MSAs, and these differences are well associated with differential changes in air pollution levels.

8. Conclusion

Air pollution and congestion are major concerns for cities around the world. There is an ongoing debate as to whether ride-hailing services such as Uber and Lyft decrease or increase air pollution and congestion. We investigate this question empirically, exploiting staggered rollout of ride-hailing entry into U.S. cities. Though ours is not the first to investigate the question, we answer it with three new approaches: (1) we construct MSA-level de facto entry dates from the Google Trends Index for both Uber and Lyft; (2) we use satellite-based data, which allow us to compare spatially delineated subareas of MSAs consistently over time; and (3) we take three complementary identification strategies, namely difference-in-differences with stratification, the fingerprint rule in Austin as a quasi-experiment, and the instrumental variable method (along with the difference-in-differences with stratification). The paper provides robust evidence that ride-hailing entry tends to reduce air pollution in large, dense cities, but has no significant effect in lower dense cities. We further support these results by another set of evidence, showing that private car use declines, with an increase in use of public transit, for large cities.

Our results are also consistent with recent studies that find mixed and heterogeneous impacts of public transportation infrastructures [Chen-Whalley (2012); Li *et al.* (2019); Gendron-Carrier *et al.* (2022)]. In a study covering 58 subway openings worldwide, for example, Gendron-Carrier *et al.* (2022) find highly heterogeneous impacts of subway openings: in 12 cities, subway openings had no effect; in 20 cities, air quality got worse; and air quality improved in 23 cities with initially high levels of air pollution. Gendron-Carrier *et al.* argue that ridership is a key driver explaining these heterogeneous impacts — meaningful air pollution reduction occurred in cities where ridership is largest. The same goes for ride-hailing — the pollution-reducing effect of ride-hailing is largest in cities where its complementarity with public transit is highest.

Our results also have important policy implications. City and urban planning practitioners around the world are increasingly concerned with the effect of ride-hailing on air pollution and congestion in urbanized areas. In response to this rising concern, Uber launched a new Uber Green service and made a $800 million fund available for Uber drivers to transit to EVs by 2025 in Canada, Europe, and the U.S. While we agree that such an effort would likely reduce air pollution *per hired ride*, its effect on carbon emissions and congestion may be ambiguous as ride-hailing may still increase hours of driving. In this context, our results suggest, in line with Agrawal and Zhao (2023), that policies that incentivise complementarity between hired rides and mass transit commuting, say, via monetary incentives for commuters or better public transportation planning, may have the double effects of increasing ridership while reducing air pollution and carbon emissions from road transportation.

References

[1] Agrawal, David R. and Weihua Zhao (2023) Taxing Uber. *Journal of Public Economics* 221,104862

[2] Anas, Alex, Robin Lindsey (2011) Reducing Urban Road Transportation Externalities: Road Pricing in Theory and in Practice. *Review of Environmental Economics and Policy* 5 (1): 66–88

[3] Anderson, Michael L. and Lucas W. Davis (2023) Uber and Traffic Fatalities. Forthcoming at *the Review of Economics and Statistics*. DOI: https://doi.org/10.1162/rest_a_01385

[4] Arnott, Richard. (1996) Taxi Travel Should Be Subsidized. *Journal of Urban Economics* 40 (3): 316-333

[5] Angrist, Joshua D., Sydnee Caldwell, Jonathan V. Hall. (2021) Uber vs. Taxi: A Driver's Eye View. *American Economic Journal: Applied Economics* 13 (3): 272-308

[6] Baker, Andrew and Larcker, David F. and Wang, Charles C. Y. (2021) How Much Should We Trust Staggered Difference-In-Differences Estimates? *Rock Center for Corporate Governance at Stanford University Working Paper* No. 246

[7] Basker, Emek (2005) Job Creation or Destruction? Labor Market Effects of Wal-Mart Expansion. *The Review of Economics and Statistics* 87 (1): 174–183

[8] Baum-Snow, Nathaniel. (2007) Did Highways Cause Suburbanization? *The Quarterly Journal of Economics* 122 (2): 775–805

[9] Berger, Thor, Chinchih Chen, Carl Benedikt Frey. (2018) Drivers of Disruption? Estimating the Uber Effect. *European Economic Review* 110: 197–210

[10] Buchholz, Nicholas (2022) Spatial Equilibrium, Search Frictions, and Dynamic Efficiency in the Taxi Industry. *The Review of Economic Studies* 89 (2): 556–591

[11] Caetano, Carolina and Brantly Callaway. (2023) Difference-in-Differences with Time-Varying Covariates in the Parallel Trends Assumption. Working paper available online at arXiv:2202.02903v2.

[12] Callaway, Brantly and Pedro H.C. Sant'Anna. (2020) Difference-in-Differences with Multiple Time Periods. *Journal of Econometrics* in press.

[13] Cape, J.N., Y.S. Tang, N. van Dijk, L. Love, M.A. Sutton, S.C.F. Palmer. (2004) Concentrations of ammonia and nitrogen dioxide at roadside verges, and their contribution to nitrogen deposition. *Environmental Pollution* 132 (3): 469-478

[14] Castillo, Juan Camilo (2023) Who Benefits from Surge Pricing? Available online at SSRN: https://ssrn.com/abstract=3245533

[15] Chen, M. Keith, Judith A. Chevalier, Peter E. Rossi, and Emily Oehlsen. (2019) The Value of Flexible Work: Evidence from Uber Drivers. *Journal of Political Economy* 127 (6): 2735-2794

[16] Chen, Yihsu, Alexander Whalley. (2012) Green Infrastructure: The Effects of Urban Rail Transit on Air Quality. *American Economic Journal: Economic Policy* 4 (1): 58–97

[17] Cohen, Peter, Robert Hahn, Jonathan Hall, Steven Levitt & Robert Metcalfe. (2016) Using Big Data to Estimate Consumer Surplus: The Case of Uber. *NBER Working Paper* 22627

[18] Correia, Sergio. (2017) Linear Models with High-Dimensional Fixed Effects: An Efficient and Feasible Estimator. Working Paper. http://scorreia.com/research/hdfe.pdf.

[19] de Chaisemartin, Clement, Xavier D'Haultfoeuille (2020) Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects. *American Economic Review*

[20] de Chaisemartin, Clement, Xavier D'Haultfoeuille (2024) Difference-in-Differences Estimators of Intertemporal Treatment Effects. Forthcoming at *the Review of Economics and Statistics* DOI: https://doi.org/10.1162/rest_a_01414

[21] Diao, M., Kong, H. & Zhao, J. Impacts of Transportation Network Companies on Urban Mobility. *Nature Sustainability* (2021). doi.org/10.1038/s41893-020-00678-z

[22] Dills, Angela K., Sean E. Mulholland. (2018) Ride-Sharing, Fatal Crashes, and Crime. *Southern Economic Journal* 84 (4): 965–991

[23] Duflo, Esther (2001) Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment. *The American Economic Review* 91 (4): 795–813.

[24] Duranton, Gilles, Matthew A. Turner. (2011) The Fundamental Law of Road Congestion: Evidence from US Cities. *American Economic Review* 101: 2616–2652

[25] Duranton, Gilles, Matthew A. Turner. (2012) Urban Growth and Transportation. *The Review of Economic Studies* 79 (4): 1407–1440

[26] Erhardt, Gregory D., Sneha Roy, Drew Cooper, Bhargava Sana, Mei Chen, Joe Castiglione. Do Transportation Network Companies Decrease or Increase Congestion? *Science Advances* 5 (5), aau2670 DOI: 10.1126/sciadv.aau2670

[27] Faber, Benjamin. (2014) Trade Integration, Market Size, and Industrialization: Evidence from China's NationalTrunk Highway System. *The Review of Economic Studies* 79 (4): 1407–1440

[28] Farber, H. (2015) Why You Can't Find A Taxi in the Rain and Other Labor Supply Lessons from Cab Drivers. *Quarterly Journal of Economics* 130 (4): 1975–2026

[29] Forbes (2019) Uber Adding To Air Pollution In Europe - Report, Nov 20, 2019, available online at www.forbes.com/sites/davekeating/2019/11/20/uber-adding-to-air-pollution-in-europereport

[30] Fréchette, Guillaume R., Alessandro Lizzeri, and Tobias Salz. (2019) Frictions in a Competitive, Regulated Market: Evidence from Taxis. *American Economic Review* 109 (8): 2954-92

[31] Gardner, John. (2022) Two-stage differences in differences. *Working Paper* 2207.05943 available at arXiv.org.

[32] Gendron-Carrier, Nicolas, Marco Gonzalez-Navarro, Stefano Polloni, Matthew A. Turner (2022) Subways and Urban Air Pollution. *American Economic Journal: Applied Economics* 14 (1): 164-96

[33] Goodman-Bacon, Andrew. (2021) Difference-in-Differences with Variation in Treatment Timing. *Journal of Econometrics* in press.

[34] Grainger, Corbett, Andrew Schreiber (2019) Discrimination in Ambient Air Pollution Monitoring? *AEA Papers and Proceedings* 109: 277-82.

[35] Gu, Yizhen, Chang Jiang, Junfu Zhang, and Ben Zou (2021) Subways and Road Congestion. *American Economic Journal: Applied Economics* 13 (2): 83–115

[36] Hall, Jonathan D., Craig Palsson, Joseph Price (2018) Is Uber A Substitute or Complement for Public Transit? *Journal of Urban Economics* 108: 36-50

[37] Hall, Jonathan V., John J. Horton, Daniel T. Knoepfle (2020) Ride-Sharing Markets Re-Equilibrate. *Working Paper*

[38] Hudson, Sally, Peter Hull, and Jack Liebersohn (2017) Interpreting Instrumented Difference-in-Differences. Unpublished manuscript available online at http://www.mit.edu/~liebers/DDIV.pdf

[39] Kim, Yeong Jae and Sarmiento, Luis. (2021) The Air Quality Effects of Uber. *RFF Working Paper Series 21-34, Resources for the Future.*

[40] Kong, Hui, Xiaohu Zhang, Jinhua Zhao (2020) How Does Ridesourcing Substitute for Public Transit? A Geospatial Perspective in Chengdu, China. *Journal of Transport Geography* 86: 102769

[41] Krishnamurthy, Chandra Kiran B. and Nicole S Ngo (2024) Do Ride-Hailing Services Worsen Freeway Congestion and Air Quality? Evidence from Ubers Entry in California. Forthcoming at *Journal of the Association of Environmental and Resource Economists*

[42] Lagos, R. (2003) An Analysis of the Market for Taxicab Rides in New York City. *International Economic Review* 44 (2): 423-434

[43] Li, Shanjun, Yanyan Liu, Avralt-Od Purevjav, Lin Yang (2019) Does Subway Expansion Improve Air Quality? *Journal of Environmental Economics and Management* 96: 213-235,

[44] Li, Ziru, Yili Hong, Zhongju Zhang. (2021) Do On-demand Ride-sharing Services Affect Traffic Congestion? Evidence from Uber Entry. Available at SSRN: https://ssrn.com/abstract=2838043

[45] Liu, Meng, Erik Brynjolfsson, and Jason Dowlatabadi. (2021) Do Digital Platforms Reduce Moral Hazard? The Case of Uber and Taxis. *NBER Working Paper* 25015

[46] Rayle, Lisa, Danielle Dai, Nelson Chan, Robert Cervero, Susan Shaheen (2016) Just A Better Taxi? A Survey-based Comparison of Taxis, Transit, and Ridesourcing Services in San Francisco. *Transport Policy* 45: 168-178

[47] Redding, Stephen J., Matthew A. Turner (2015) Transportation Costs and the Spatial Organization of Economic Activity. Edited by Gilles Duranton, J. Vernon Henderson, William C. Strange, *Handbook of Regional and Urban Economics* 5, 1339-1398

[48] Neumark, David, Junfu Zhang, Stephen Ciccarella (2008) The Effects of Wal-Mart on Local Labor Markets. *Journal of Urban Economics* 63 (2): 405-430

[49] Second Measure (2021) Uber vs. Lyft: Who's Tops in the Battle of U.S. Rideshare Companies. An online article posted by Liyin Yeo, March 16, 2021, available at https://secondmeasure.com/datapoints/rideshare-industry-overview

[50] Słoczyński, Tymon. (2022) Interpreting OLS Estimands When Treatment Effects Are Heterogeneous: Smaller Groups Get Larger Weights. *Review of Economics and Statistics* 104 (3): 501–509.

[51] Su, Jason G., Michael Jerrett, Bernardo Beckerman, Michelle Wilhelm, Jo Kay Ghosh, Beate Ritz (2009) Predicting traffic-related air pollution in Los Angeles using a distance decay regression selection strategy. *Environmental Research* 109 (6): 657-670

[52] Sun, Liyang, Sarah Abraham (2020) Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects. *Journal of Econometrics*

[53] Tarduno, Matthew (2021) The congestion costs of Uber and Lyft. *Journal of Urban Economics* 122, 103318

[54] Union of Concerned Scientists (2020) Ride-Hailing's Climate Risks: Steering a Growing Industry toward a Clean Transportation Future. An Independent Report: Cambridge, MA. February 2020

[55] U.S. Environmental Protection Agency (EPA). (2008) *Risk and Exposure Assessment to Support the Review of the NO2 Primary National Ambient Air Quality Standard.* Office of Air Quality Planning and Standards

[56] Research Triangle Park, North Carolina. # EPA-452/R-08-008a

[57] Ward, Jacob W., Jeremy J. Michalek, Ines L. Azevedo, Constantine Samaras, Pedro Ferreira (2019) Effects of On-demand Ridesourcing on Vehicle Ownership, Fuel Consumption, Vehicle Miles Traveled, and Emissions Per Capita in U.S. States. *Transportation Research Part C: Emerging Technologies* 108: 289-301

[58] Wooldridge, Jeffrey M. (2021) Two-Way Fixed Effects, the Two-Way Mundlak Regression, and Event Study Estimators. *Working paper*

[59] Zou, Eric Yongchen (2021) Unwatched Pollution: The Effect of Intermittent Monitoring on Air Quality. *American Economic Review* 111 (7): 2101-26.

Figure 1. Complementarity Effect of Ride-hailing Entry on Daily Transport Choices

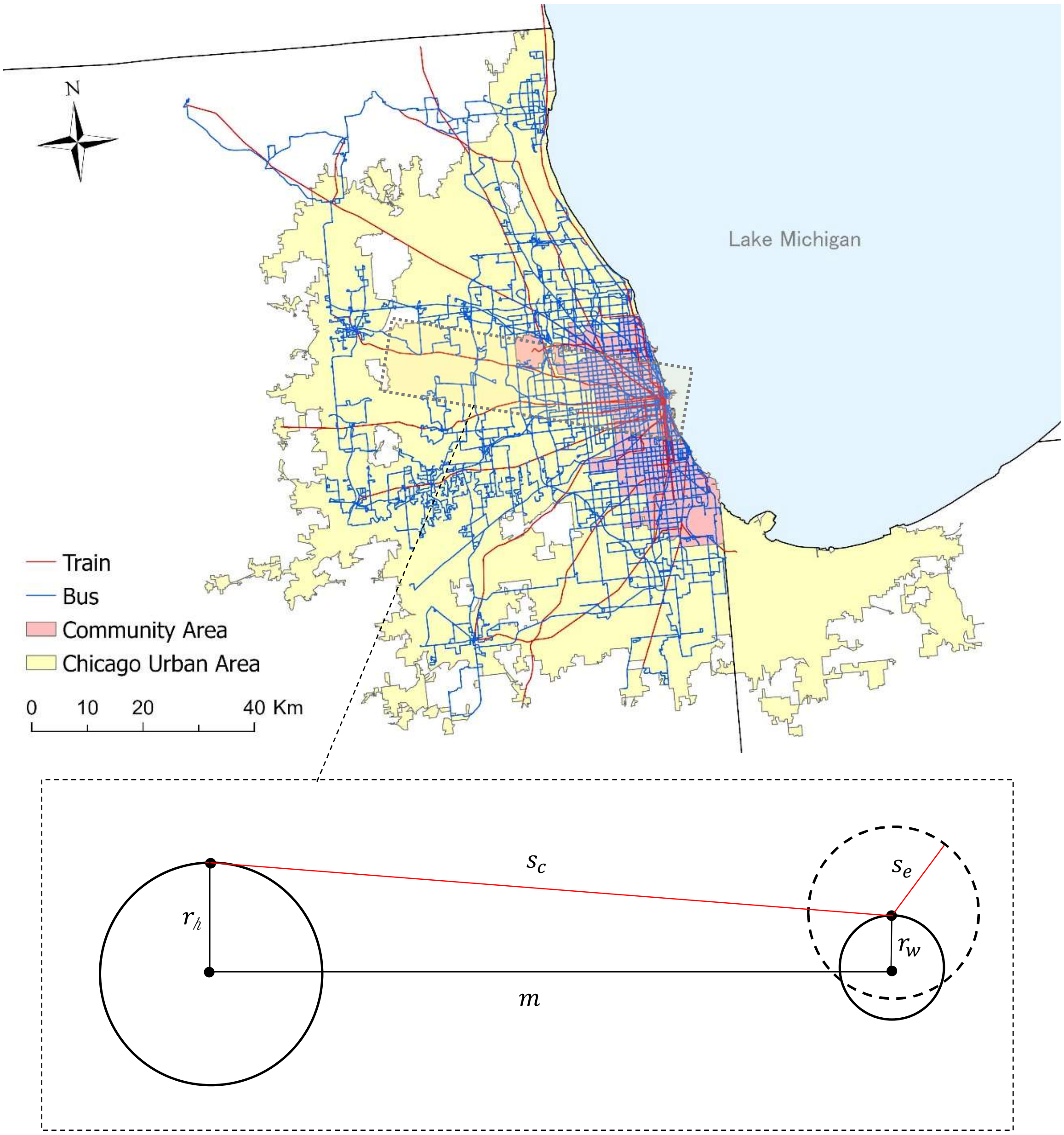
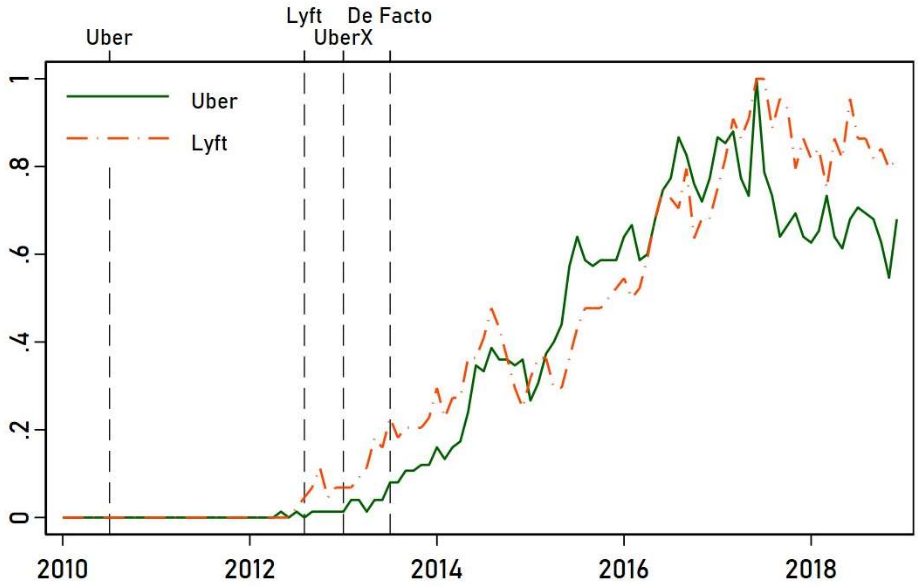


Figure 2. Google Trends as Measure of Uber/Lyft Market Presence

(a) San Francisco, CA



(b) Austin, TX

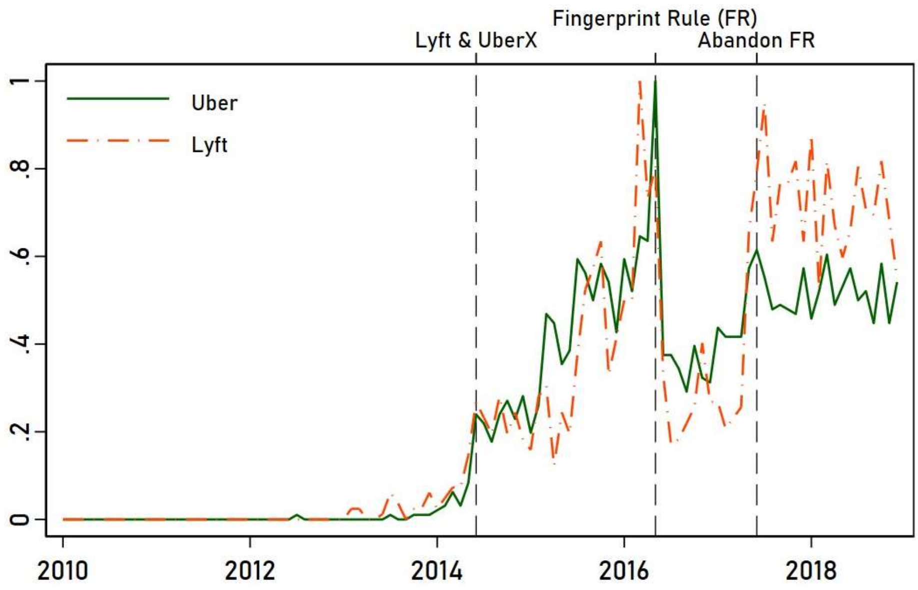
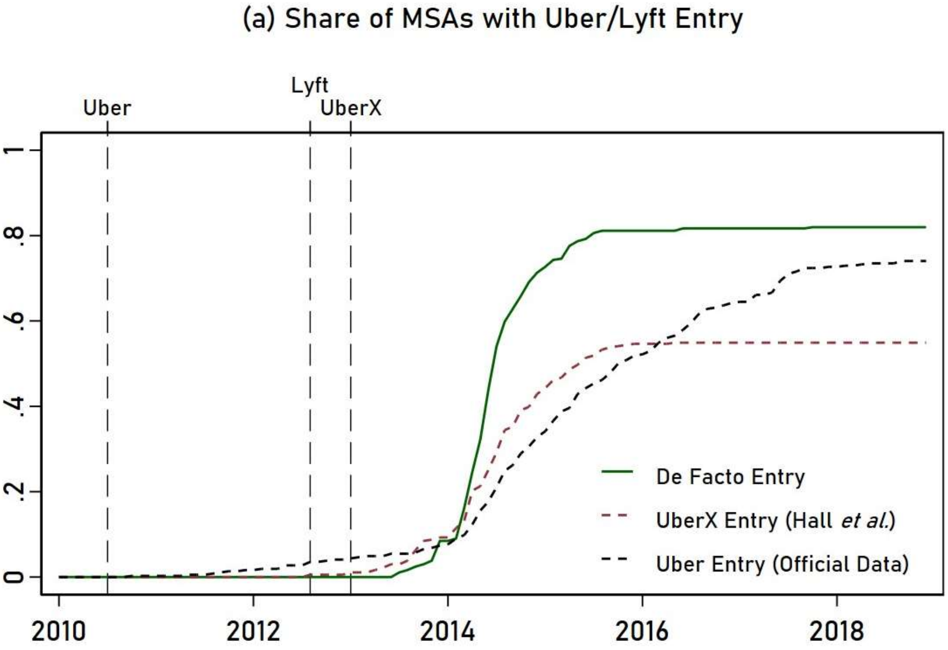
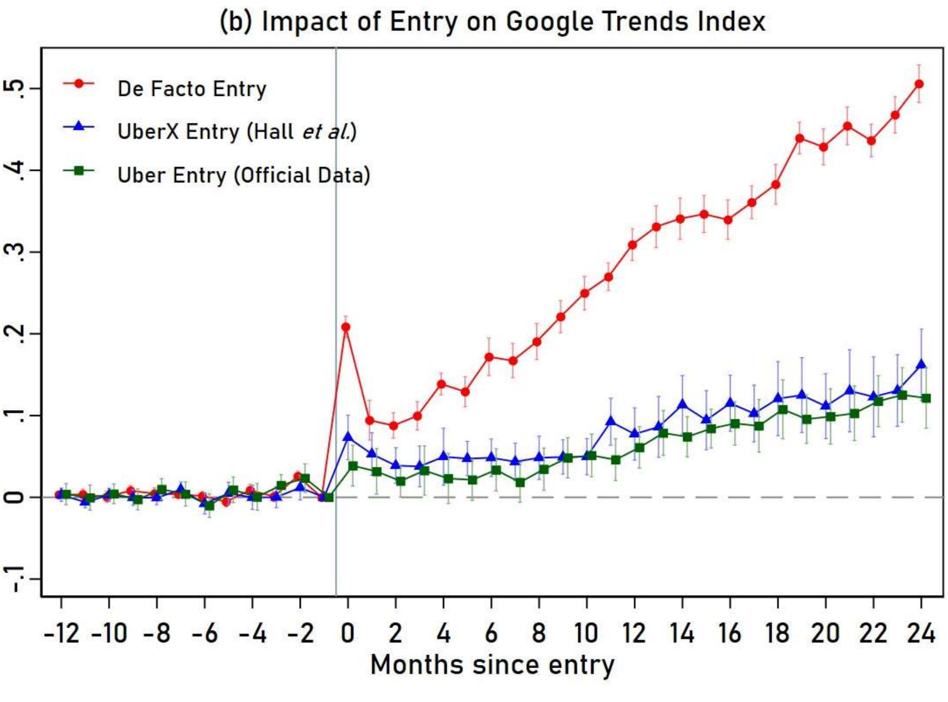


Figure 3. De Facto versus Reported Entry

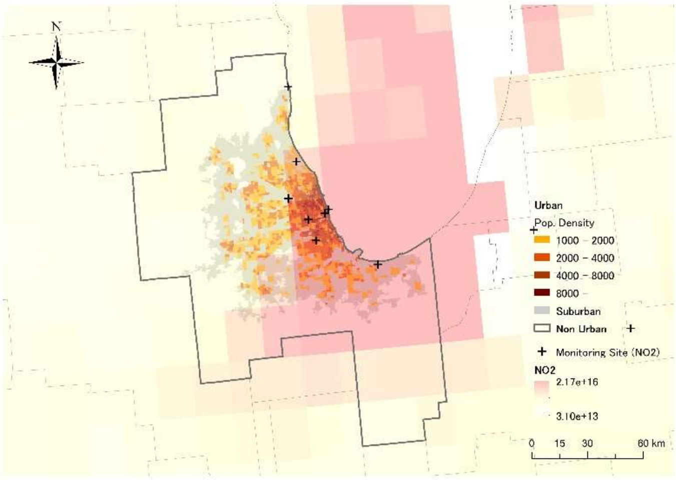




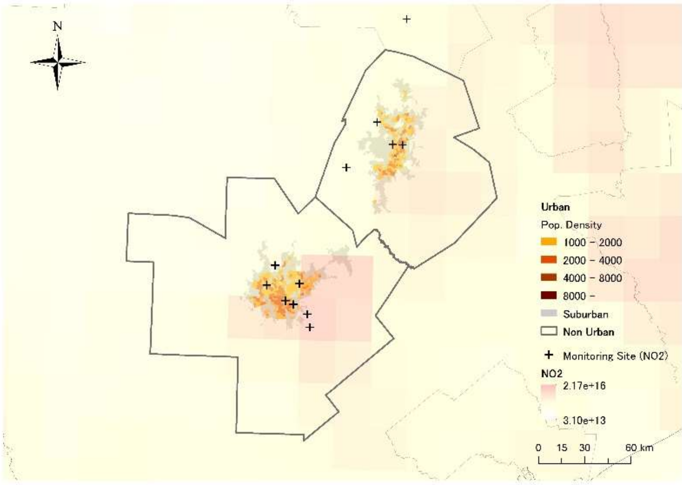
Note: The top panel report the cumulative distribution of ride-hailing entry using three alternative measures of entry: our measure based on Google Trends indices, UberX entry dates from Hall *et al.* (2018), and official Uber entry dates. The bottom panel report the event study estimates regressing entry timings on the composite Google Trends index using alternative entry measures. Whiskers represent the 95% confidence intervals using robust standard errors clustered at the MSA level.

Figure 4. Satellite-based versus Monitoring Data on NO_2 Concentration
Monthly Average in January 2014

A. Chicago, IL



B. Austin and San Antonio, TX



C. San Francisco, CA

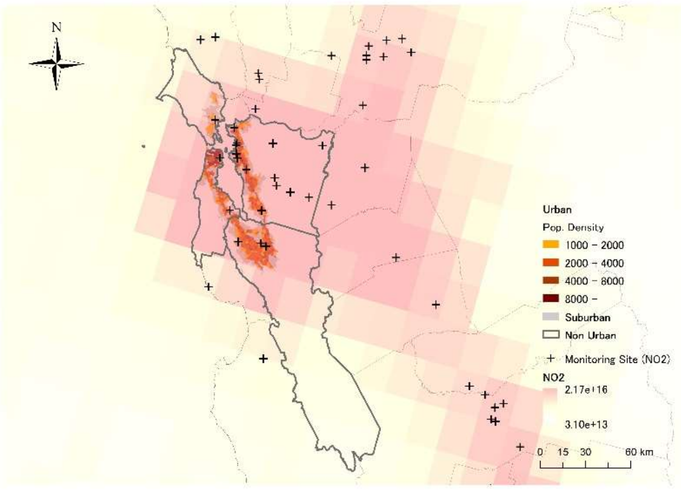
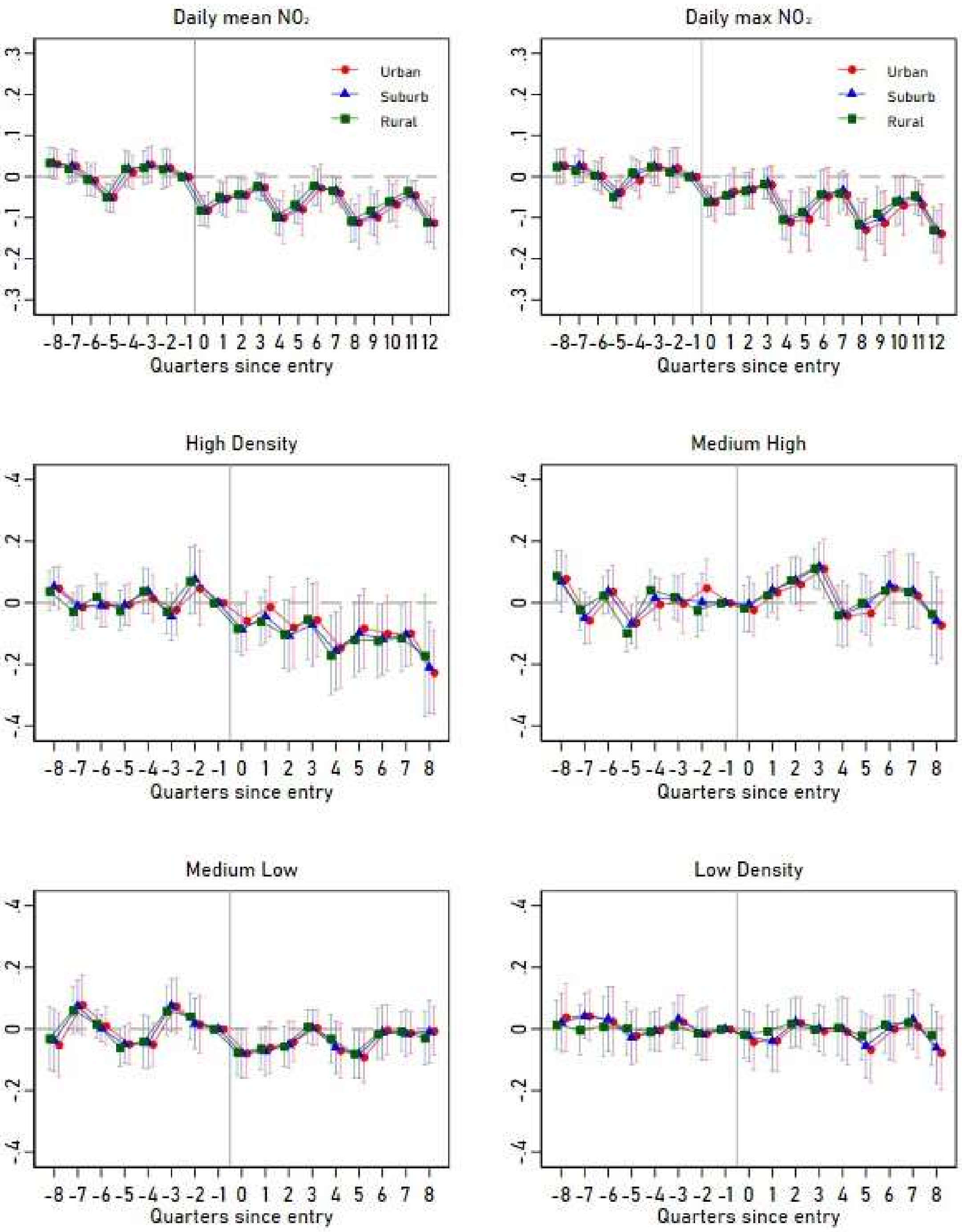
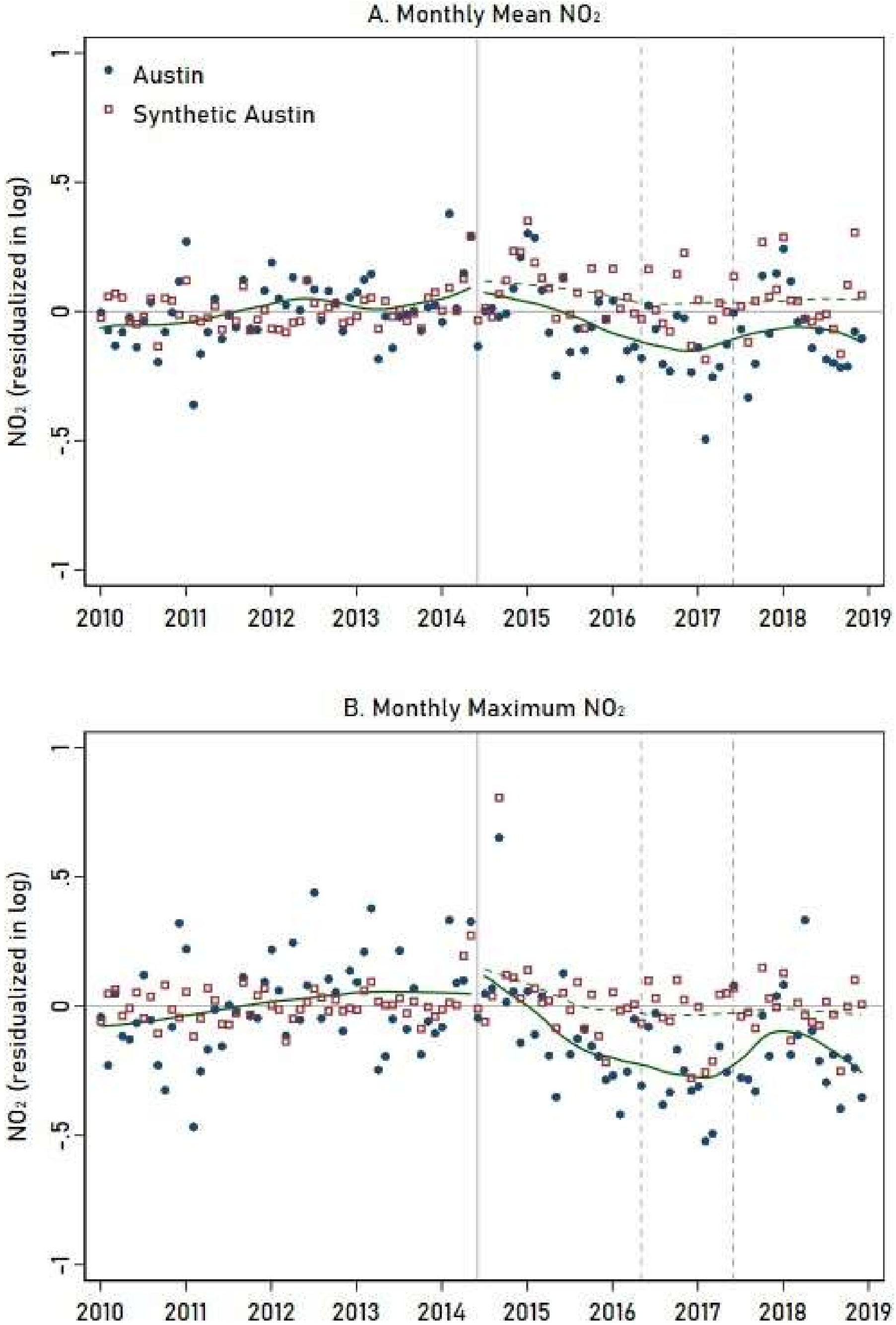


Figure 5. Event Study Estimates of
the Effect of Entry on NO_2 Concentrations



Note: The two graphs on the top panel report the event study estimates from the dCDH estimator on the pooled sample. The other four graphs plot the event study estimates on each population density quartile using monthly maximum NO_2 concentrations. Whiskers represent the 95% confidence intervals using robust standard errors clustered at the MSA level.

Figure 6. Synthetic vs. Observed NO_2 Trends in Austin, Texas



Note: Circles (navy) are the predicted residuals from the regression of monthly NO_2 concentrations on all covariates and fixed effects using only pre-treatment data. Squares (maroon) are the counterfactual NO_2 concentration residuals based on the synthetic control method using the 12-month lags as covariates. The solid lines plot the local polynomial smoothing of the actual residuals. The dashed lines plot the local polynomial smoothing of the counterfactual residuals.

Table 1. Pre-treatment Conditions by Treatment Status

	No Entry		Entry in 2013		Entry in 2014		After 2015		All	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
N. of MSAs	62		30		217		39		348	
(Share of full sample)	(0.18)		(0.09)		(0.62)		(0.11)		(1.00)	
Pop. density	117	(79.68)	385	(538.30)	335	(339.83)	228	(131.77)	288	(327.03)
Median household income	51,635	(7,592)	60,592	(13,991)	57,361	(10,759)	55,307	(8,200)	56,389	(10,592)
Median age	41.3	(3.906)	39.8	(3.842)	41.1	(4.018)	40.7	(3.900)	41.0	(3.973)
Share of white	0.86	(0.109)	0.80	(0.132)	0.82	(0.112)	0.85	(0.110)	0.83	(0.114)
Share of college degree	0.16	(0.044)	0.20	(0.062)	0.19	(0.059)	0.17	(0.042)	0.19	(0.056)
Manf. employment share	0.10	(0.057)	0.11	(0.045)	0.12	(0.070)	0.12	(0.061)	0.11	(0.065)
Share of private car commuters	0.91	(0.036)	0.89	(0.045)	0.90	(0.043)	0.91	(0.031)	0.90	(0.041)
Share of pub. transit commuters	0.01	(0.007)	0.02	(0.026)	0.02	(0.026)	0.01	(0.010)	0.01	(0.022)
NO_2 concentrations (mean)										
EPA (ppb)	6.54	(5.094)	12.34	(3.155)	9.52	(3.978)	7.38	(0.850)	9.57	(4.087)
Satellite (log of molec./ cm^2)	35.17	(0.498)	35.46	(0.487)	35.54	(0.438)	35.47	(0.362)	35.46	(0.465)
Share of EPA sample	0.08		0.27		0.22		0.05		0.18	

Note: All statistics are as of year 2010.

Table 2. TWFE Estimates of Effect of Ride-hailing Entry on Ambient Air Quality

A. Monthly Mean NO_2 Concentrations

	Mean NO_2 Concentration								
	Urban Area			Suburban Area			Non-urban Area		
Entry									
× 1st QuartileMSAs	0.000	0.007	0.007	0.002	0.008	0.008	0.025	0.033 *	0.033
	(0.015)	(0.016)	(0.016)	(0.015)	(0.015)	(0.016)	(0.017)	(0.017)	(0.018)
× 2nd Quartile MSAs	-0.018	-0.013	-0.01	-0.018	-0.012	-0.01	-0.016	-0.01	-0.006
	(0.012)	(0.011)	(0.011)	(0.011)	(0.010)	(0.011)	(0.010)	(0.009)	(0.010)
× 3rd Quartile MSAs	-0.013	-0.008	-0.005	-0.014	-0.009	-0.006	-0.014	-0.009	-0.005
	(0.014)	(0.015)	(0.014)	(0.013)	(0.014)	(0.013)	(0.012)	(0.013)	(0.012)
× 4th Quartile MSAs	-0.044	-0.038 **	-0.034 **	-0.042 **	-0.036 **	-0.033 **	-0.041 **	-0.036 **	-0.032 **
	(0.014)	(0.014)	(0.013)	(0.014)	(0.013)	(0.013)	(0.013)	(0.013)	(0.012)
Controls included:									
Year/Month FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
MSA FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Climate controls		✓	✓		✓	✓		✓	✓
Other controls			✓			✓			✓
Obs.	38,911	35,023	35,023	39,309	35,421	35,421	39,254	35,367	35,367
Adj. R^2	0.759	0.757	0.757	0.765	0.762	0.762	0.781	0.778	0.778

Note: In parenthesis are robust standard errors clustered at the MSA level.

B. Monthly Maximum NO_2 Concentrations

	Max NO_2 Concentration								
	Urban Area			Suburban Area			Non-urban Area		
Entry									
× 1st QuartileMSAs	-0.011	-0.007	-0.007	-0.009	-0.005	-0.006	0.012	0.017	0.018
	(0.017)	(0.018)	(0.018)	(0.016)	(0.017)	(0.017)	(0.014)	(0.014)	(0.015)
× 2nd Quartile MSAs	-0.027	-0.025	-0.023	-0.028	-0.027	-0.024	-0.026	-0.024	-0.020
	(0.019)	(0.019)	(0.019)	(0.020)	(0.019)	(0.019)	(0.015)	(0.014)	(0.016)
× 3rd Quartile MSAs	-0.030 *	-0.024	-0.021	-0.031 **	-0.025 *	-0.022	-0.032 **	-0.027 *	-0.021
	(0.013)	(0.014)	(0.013)	(0.013)	(0.014)	(0.013)	(0.011)	(0.012)	(0.012)
× 4th Quartile MSAs	-0.059 **	-0.054 **	-0.050 **	-0.058 **	-0.054 **	-0.049 **	-0.059 **	-0.055 **	-0.050 **
	(0.021)	(0.020)	(0.018)	(0.020)	(0.019)	(0.018)	(0.019)	(0.017)	(0.016)
Controls included:									
Year/Month FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
MSA FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Climate controls		✓	✓		✓	✓		✓	✓
Other controls			✓			✓			✓
Obs.	38,911	35,023	35,023	39,309	35,421	35,421	39,254	35,367	35,367
Adj. R^2	0.662	0.66	0.661	0.679	0.677	0.677	0.712	0.709	0.709

Note: In parenthesis are robust standard errors clustered at the MSA level.

Table 3. TWFE Estimates of Effect of Ride-hailing Entry on Commuting Mode Choice

A. Primary Commuting Mode

	Commuting Mode to Work								
	Private Car			Public Transit			Other Mode		
Entry									
× 1st QuartileMSAs	0.002	0.003 *	0.003 *	0.000	-0.001	-0.001	-0.001	-0.002	-0.002
	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
× 2nd Quartile MSAs	-0.001	0.000	0.001	-0.001	-0.001	-0.001	0.002	0.000	0.000
	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
× 3rd Quartile MSAs	-0.006 ***	-0.005 ***	-0.004 ***	0.000	0.000	-0.001	0.006 ***	0.005 ***	0.004 ***
	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
× 4th Quartile MSAs	-0.014 ***	-0.014 ***	-0.013 ***	0.005 *	0.005 *	0.004 *	0.008 ***	0.007 ***	0.007 ***
	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
Controls included:									
Year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
MSA FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Demog. Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Climate controls		✓	✓		✓	✓		✓	✓
Other controls			✓			✓			✓
Obs.	18,936,750	17,872,388	17,872,388	18,936,750	17,872,388	17,872,388	18,936,750	17,872,388	17,872,388
Adj. R^2	0.061	0.063	0.063	0.115	0.116	0.116	0.007	0.007	0.007

Note: In parenthesis are robust standard errors clustered at the MSA level. Repeated cross-section, household-level commuting mode data from American Community Survey (ACS). Demographic controls include household-level covariates.

B. Other Commuting Mode: Subcategories

	Other Commuting Mode to Work								
	Walking			Work at Home			Other Mode		
Entry									
× 1st QuartileMSAs	-0.001	-0.001	-0.001	0.000	-0.001	-0.001	-0.001	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
× 2nd Quartile MSAs	-0.001	-0.001 *	-0.001 *	0.002 **	0.001	0.001	0.000	0.000	0.000
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
× 3rd Quartile MSAs	0.000	0.000	0.000	0.006 ***	0.005 ***	0.005 ***	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
× 4th Quartile MSAs	0.002 ***	0.001 ***	0.001 ***	0.006 ***	0.005 ***	0.005 ***	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
Controls included:									
Year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
MSA FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Demog. Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Climate controls		✓	✓		✓	✓		✓	✓
Other controls			✓			✓			✓
Obs.	18,936,750	17,872,388	17,872,388	18,936,750	17,872,388	17,872,388	18,936,750	17,872,388	17,872,388
Adj. R^2	0.007	0.007	0.007	0.010	0.010	0.010	0.002	0.002	0.002

Note: In parenthesis are robust standard errors clustered at the MSA level. Repeated cross-section, household-level commuting mode data from American Community Survey (ACS). Demographic controls include household-level covariates.

Table 4. TWFE Estimates of Effect of Ride-hailing Entry on Ambient Air Quality
Texas MSAs Only

A. Monthly Mean NO_2 Concentrations

	Mean NO_2 Concentration								
	Urban Area			Suburban Area			Non-urban Area		
Entry									
× 1st QuartileMSAs	-0.006 (0.016)	-0.003 (0.017)	-0.011 (0.018)	-0.001 (0.016)	0.002 (0.016)	-0.005 (0.017)	-0.003 (0.014)	-0.001 (0.014)	-0.005 (0.016)
× 2nd Quartile MSAs	0.021 (0.015)	0.026 (0.016)	0.023 (0.016)	0.021 (0.014)	0.026 * (0.014)	0.024 (0.014)	0.024 * (0.013)	0.027 ** (0.013)	0.025 * (0.013)
× 3rd Quartile MSAs	-0.005 (0.013)	0.000 (0.014)	-0.003 (0.012)	0.003 (0.011)	0.008 (0.012)	0.006 (0.010)	0.004 (0.012)	0.007 (0.013)	0.004 (0.013)
× 4th Quartile MSAs	-0.029 *** (0.010)	-0.026 ** (0.010)	-0.018 ** (0.008)	-0.023 ** (0.009)	-0.020 ** (0.009)	-0.013 * (0.007)	-0.008 (0.009)	-0.006 (0.009)	-0.009 (0.007)
Fingerprint Rule	-0.079 *** (0.007)	-0.077 *** (0.007)	-0.072 *** (0.010)	-0.069 *** (0.007)	-0.068 *** (0.006)	-0.064 *** (0.010)	-0.066 *** (0.006)	-0.065 *** (0.006)	-0.065 *** (0.009)
Controls included:									
Year/Month FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
MSA FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Climate controls		✓	✓		✓	✓		✓	✓
Other controls			✓			✓			✓
Obs.	2,016	2,016	2,016	2,016	2,016	2,016	2,016	2,016	2,016
Adj. R^2	0.767	0.772	0.772	0.764	0.769	0.769	0.762	0.767	0.767

Note: In parenthesis are robust standard errors clustered at the MSA level.

B. Monthly Maximum NO_2 Concentrations

	Max NO_2 Concentration								
	Urban Area			Suburban Area			Non-urban Area		
Entry									
× 1st QuartileMSAs	-0.019 (0.020)	-0.015 (0.019)	-0.021 (0.019)	-0.014 (0.018)	-0.009 (0.017)	-0.013 (0.017)	-0.007 (0.012)	-0.003 (0.012)	-0.006 (0.013)
× 2nd Quartile MSAs	0.031 *** (0.011)	0.037 *** (0.011)	0.035 *** (0.010)	0.025 ** (0.012)	0.032 ** (0.012)	0.031 *** (0.011)	0.018 * (0.010)	0.024 ** (0.010)	0.023 ** (0.010)
× 3rd Quartile MSAs	-0.048 *** (0.010)	-0.043 *** (0.011)	-0.040 *** (0.009)	-0.033 * (0.017)	-0.026 (0.018)	-0.024 (0.015)	-0.020 * (0.010)	-0.014 (0.011)	-0.015 (0.011)
× 4th Quartile MSAs	-0.079 *** (0.019)	-0.075 *** (0.020)	-0.054 *** (0.013)	-0.075 *** (0.016)	-0.070 *** (0.016)	-0.054 *** (0.012)	-0.043 *** (0.011)	-0.039 *** (0.011)	-0.035 *** (0.010)
Fingerprint Rule	-0.123 *** (0.010)	-0.118 *** (0.010)	-0.123 *** (0.010)	-0.103 *** (0.009)	-0.099 *** (0.008)	-0.102 *** (0.010)	-0.098 *** (0.009)	-0.094 *** (0.009)	-0.094 *** (0.009)
Controls included:									
Year/Month FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
MSA FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Climate controls		✓	✓		✓	✓		✓	✓
Other controls			✓			✓			✓
Obs.	2,016	2,016	2,016	2,016	2,016	2,016	2,016	2,016	2,016
Adj. R^2	0.689	0.692	0.692	0.704	0.708	0.707	0.723	0.727	0.726

Note: In parenthesis are robust standard errors clustered at the MSA level.

Table 5. TWFE Estimates of Effect of Ride-hailing Entry on Commuting Mode Choice
Texas MSAs Only

A. Primary Commuting Mode

	Commuting Mode to Work								
	Private Car			Public Transit			Other Mode		
Entry									
× 1st QuartileMSAs	0.000	0.001	0.001	-0.001	-0.001	0.000	0.000	0.000	-0.001
	(0.005)	(0.004)	(0.004)	(0.002)	(0.002)	(0.002)	(0.004)	(0.003)	(0.003)
× 2nd Quartile MSAs	0.001	-0.002	-0.002	0.001	0.001	0.001	-0.003	0.000	0.000
	(0.003)	(0.003)	(0.003)	(0.001)	(0.001)	(0.001)	(0.003)	(0.003)	(0.003)
× 3rd Quartile MSAs	0.002	-0.003	-0.003	-0.002 ***	-0.002 ***	-0.002 ***	0.001	0.005	0.005
	(0.003)	(0.004)	(0.004)	(0.000)	(0.000)	(0.000)	(0.003)	(0.004)	(0.004)
× 4th Quartile MSAs	-0.007 **	-0.006 **	-0.006 **	-0.001 **	-0.001 *	-0.001 **	0.007 **	0.007 ***	0.007 ***
	(0.003)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.003)	(0.002)	(0.002)
Fingerprint Rule	-0.010 ***	-0.010 ***	-0.011 ***	0.004 ***	0.004 ***	0.005 ***	0.006 ***	0.006 ***	0.006 ***
	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Controls included:									
Year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
MSA FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Demog. controls		✓	✓		✓	✓		✓	✓
Other controls			✓			✓			✓
Obs.	968,841	968,841	968,841	968,841	968,841	968,841	968,841	968,841	968,841
Adj. R^2	0.007	0.007	0.007	0.007	0.007	0.007	0.007	0.007	0.007

Note: In parenthesis are robust standard errors clustered at the MSA level. Climate and other time-varying controls that vary only at the state level are excluded from these regressions.

B. Other Commuting Mode: Subcategories

	Other Commuting Mode to Work								
	Walking			Work at Home			Other Mode		
Entry									
× 1st QuartileMSAs	0.003	0.002	0.002	-0.003	-0.002	-0.003	0.000	0.000	0.000
	(0.004)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
× 2nd Quartile MSAs	-0.001	0.003 **	0.003 **	0.002	0.001	0.001	-0.003 **	-0.004 **	-0.004 **
	(0.003)	(0.001)	(0.001)	(0.003)	(0.003)	(0.002)	(0.001)	(0.002)	(0.002)
× 3rd Quartile MSAs	0.000	0.001	0.001	0.002	0.006 ***	0.005 **	-0.002	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
× 4th Quartile MSAs	0.002 **	0.002 **	0.002 **	0.007 ***	0.006 ***	0.007 ***	-0.002	-0.002	-0.002
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
Fingerprint Rule	-0.004 ***	-0.004 ***	-0.004 ***	0.012 ***	0.012 ***	0.012 ***	-0.003 ***	-0.002 ***	-0.002 ***
	(0.001)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)
Controls included:									
Year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
MSA FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Demog. controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Other controls			✓			✓			✓
Obs.	968,841	968,841	968,841	968,841	968,841	968,841	968,841	968,841	968,841
Adj. R^2	0.002	0.002	0.002	0.014	0.014	0.014	0.003	0.003	0.003

Note: In parenthesis are robust standard errors clustered at the MSA level. Climate and other time-varying controls that vary only at the state level are excluded from these regressions.

Table 6. DD-IV Estimates of Effect of Ride-hailing Entry on Ambient Air Quality

	Population Density			
	Lowest MSAs	**2nd MSAs**	**3rd MSAs**	**Highest MSAs**
Mean NO_2				
Urban area	0.121	-0.119	0.029	-0.104
	(0.108)	(0.204)	(0.057)	(0.081)
Suburban area	0.097	-0.154	0.016	-0.066
	(0.108)	(0.256)	(0.056)	(0.078)
Non-urban area	0.062	-0.109	-0.012	-0.057
	(0.111)	(0.250)	(0.054)	(0.075)
Obs.	8,460	8,601	8,762	8,936
Weak IV stat.	71.49	16.09	177.53	120.90
Hansen's J stat.	0.812	1.177	5.845	0.053
(p-values)	0.368	0.555	0.054	0.974

Note: In parenthesis are robust standard errors clustered at the MSA level. We use the full set of controls for all regressions. Weak identification/overidentification statistics are reported for urban area regressions only. Regressions are separately estimated for each population density strata.

Table 7. DD-IV Estimates of Effect of Ride-hailing Entry on Commuting Mode Choice

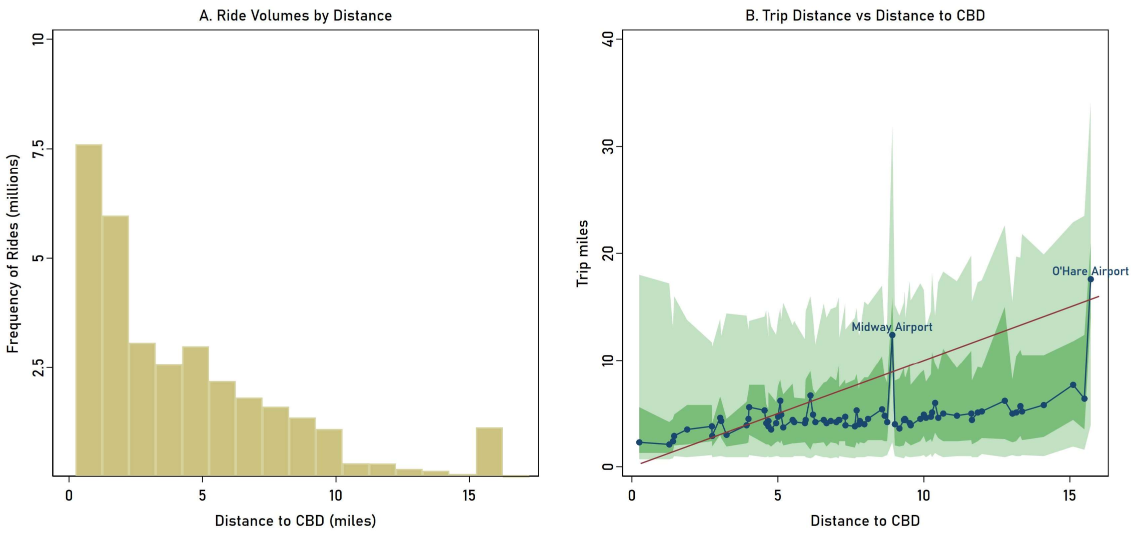
	Population Density			
	Lowest MSAs	**2nd MSAs**	**3rd MSAs**	**Highest MSAs**
Commuting Mode to Work				
Private Car	-0.001	0.013	-0.013	-0.061 *
	(0.008)	(0.024)	(0.009)	(0.036)
Public Transit	-0.001	-0.013	0.002	0.050
	(0.002)	(0.011)	(0.003)	(0.036)
Other Mode	0.002	0.001	0.009	0.007
	(0.007)	(0.016)	(0.008)	(0.012)
Obs.	1,624,938	2,066,010	3,626,698	10,464,084
Weak IV stat.	53,298	16,487	156,026	184,734
Hansen's J stat.	0.981	1.324	2.188	3.137
(p-values)	0.322	0.516	0.335	0.208

Note: In parenthesis are robust standard errors clustered at the MSA level. We use the full set of controls for all regressions. Weak identification/overidentification statistics are reported for urban area regressions only. Regressions are separately estimated for each population density strata.

Appendix A

The City of Chicago provides detailed ride-hailing trip-level data since November 2018 at data.cityofchicago.org. The data record all trips served by transportation network companies that took place within the community area boundaries. The trip data exclude taxi trips. We use the data to calculate the travel distance of each trip as well as the distance between each trip's pickup location to the centroid of the central business district (CBD) in Chicago. **Figure A1** plots (A) the frequency distribution of ride-hailing trips during the weekday rush hours in 2019 (before the Covid-19 pandemic) and (B) the distribution of travel distance against the distance to the CBD. The rush hours are defined as 6-9 am in the morning and 16-19 pm in the evening. These rush hours are determined from the observed peaks in the number of ride-hailing rides during the weekdays, and thus, may not necessarily match the rush hours in other transportation modes such as freeway traffic and public transit. From panel (A), we see a large fraction of ride-hailing rides occur outside the CBD, although the frequency of rides tends to decrease with the distance to CBD. Panel (B) shows the centiles of trip distance originating from each community area. The straight line in red is the 45-degree line. The idea here is that if a resident takes the ride to commute from the pickup location all the way to CBD, then the trip distance (on the vertical axis) should be greater than the distance to CBD (on the horizontal axis) so that it should lie above the 45-degree line. For example, the figure indicates that more than 75% of ride-hailing trips originating from O'Hare Airport travel lie above the 45-degree line, implying that customers on these trips are likely to be traveling to CBD. The figure demonstrates that roughly 75% of weekday rush-hour rides that originate from 5 miles away or more from CBD do not directly travel to CBD.

Figure A1. Geographic Distribution of Ride-hailing Trips in Chicago
Morning and Evening Rush Hours, Weekdays, 2019



Appendix B

We run the following OLS and Tobit regressions to examine the correlation between entry timing/location and MSA-level socioeconomic variables as of year 2010. For the Tobit regression, we assume the entry dates are truncated at December 2020 for no-entry MSAs.

$$EntryDate_c = \alpha + X'_{c,2010}\beta + \epsilon$$

Some socioeconomic variables are missing for three of the 348 MSAs used in our study, and hence are omitted from the regressions.

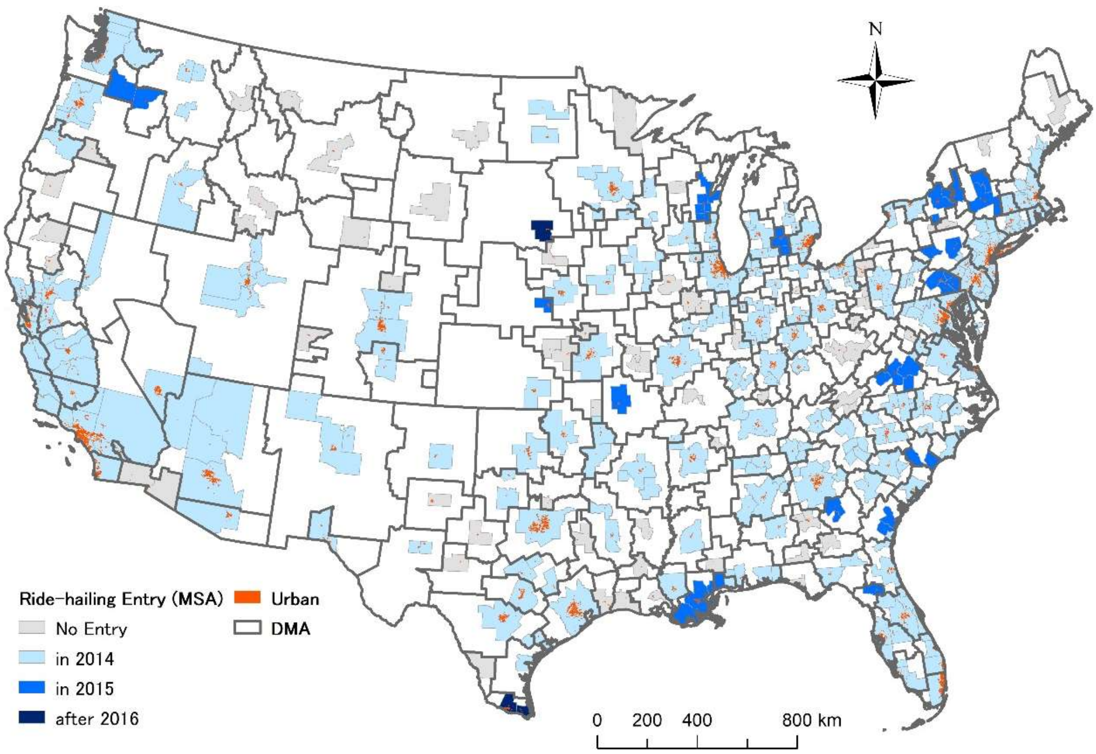
Table A1. Association between entry timing and Pre-treatment Covariates.

	Entry timing					
	OLS			Tobit		
Pop_{2010}	-3.246	(2.06)		-3.661	(1.98)	*
$Pop.\ density_{2010}$	-9.850	(2.51)	***	-11.893	(3.27)	***
$Median\ age_{2010}$	-0.258	(0.45)		-0.294	(0.58)	
$Median\ income_{2010}$	-28.261	(18.04)		-32.445	(19.63)	*
$College_{2010}$	-5.056	(46.48)		-7.377	(51.65)	
$Non\text{-}white_{2010}$	-20.169	(15.83)		-23.570	(21.79)	
$Manuf.\ share_{2010}$	-61.602	(27.70)	**	-73.591	(31.09)	**
$Poverty_{2010}$	-14.184	(100.44)		-18.643	(120.08)	
$Car\ commute_{2010}$	47.177	(68.17)		49.750	(81.96)	
$Pub.\ transit_{2010}$	200.934	(120.13)	*	234.579	(111.73)	*
N	345			345		
R^2	0.1945					

Appendix C

The following map illustrates the entry timing and location using our *de facto* entry measure. The map also shows the urban areas used to calculate urban-area NO_2 concentrations in our main analysis.

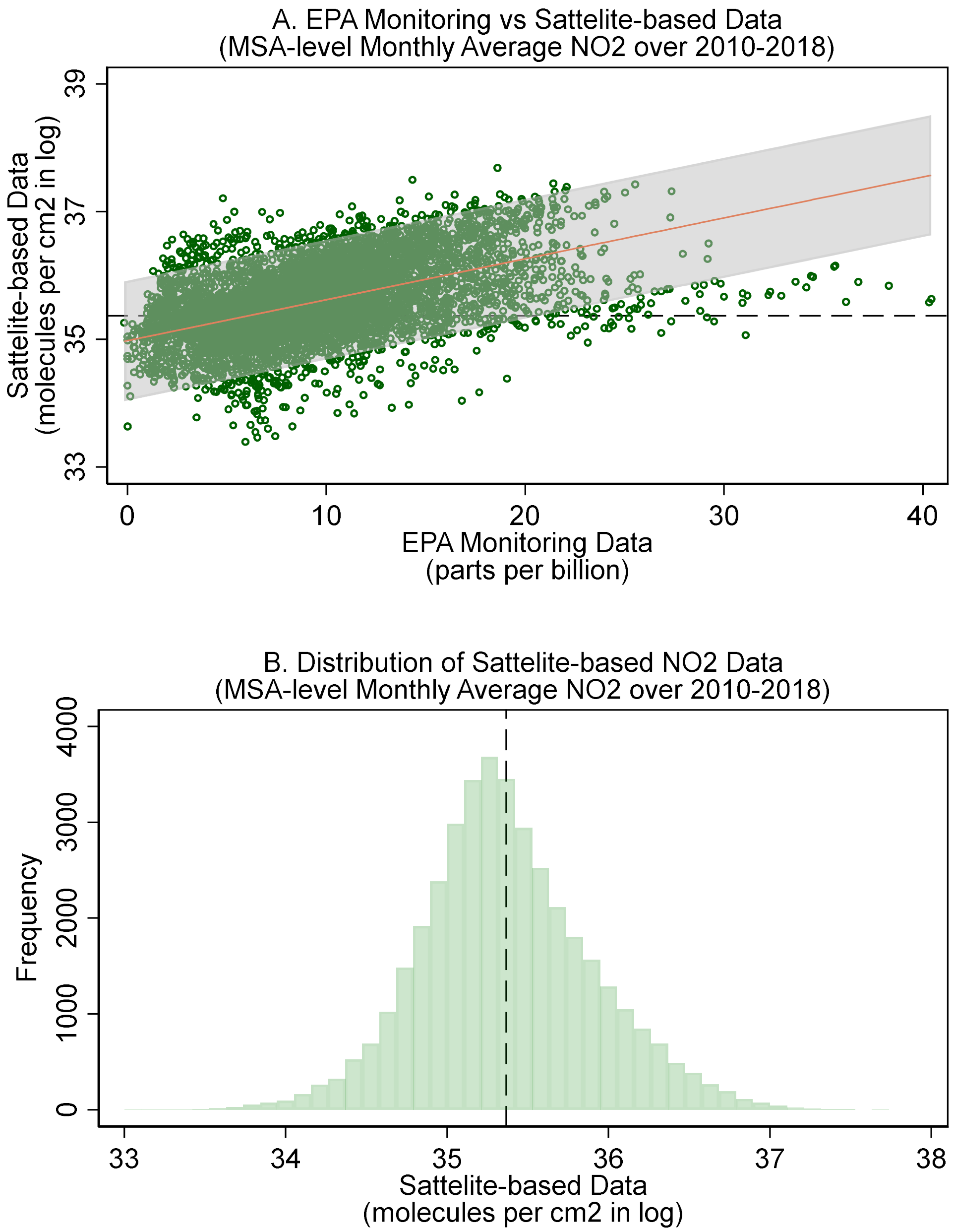
Figure A2. Geographic Distribution of Ride-haling Entry and Urban Area Boundaries



Appendix D

The following figure examines the relationship between EPA monitoring data versus NASA's satellite-based data. Although there is a clear (statistically significant) relationship between the two, there is also substantial variation in the satellite-based NO_2 concentrations for each level of EPA NO_2 concentrations. There is also an indication that EPA data might be missing for high concentration episodes, as in Zou (2021) and Grainger (2021).

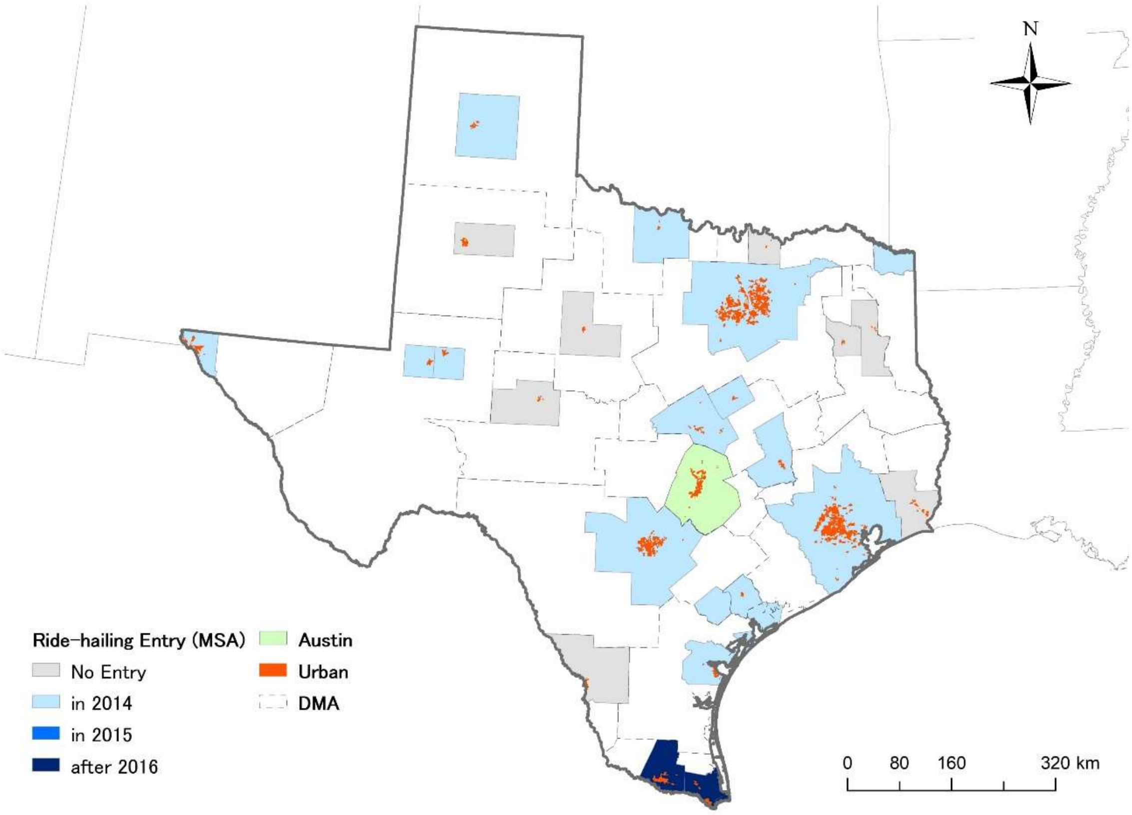
Figure A3. Satellite-based versus EPA Monitoring Data



Appendix E

The following map illustrates the entry timing and location for Texas, in a manner similar to Figure A1.

Figure A4. Geographic Distribution of Ride-haling Entry in Texas



Appendix F. Geography-based Instruments

This appendix explains how we construct our geography-based instruments.

For the highway construction plan as of 1947, we use a digital image of the 1947 highway plan from Duranton and Turner (2011; 2012) and covert it into a GIS map so that the two maps' coordinates match each other. Figure A5 is a map showing which MSAs were planned to connect on the 1947 highway plan. As Duranton and Turner note, many interstate highways were built subsequently based on this highway plan.

For the railroad network as of 1870, we use GIS shapefiles from Katherine Walter's GIS Railroads and the Making of Modern America Project at the University of Nebraska--Lincoln. Figure A6 displays historical railroad maps between 1840 and 1970 using this data. As Duranton and Turner (2011; 2012) argue, many of the railroads from this period are abandoned and turned into roads, and hence, are good predictors of MSA-level road networks and economic activities today.

To construct the Euclidean spanning tree network, we start with the cities that existed as of 1860 and had a population of 10,000 or more. We then follow the Kruskal's algorithm to compute the minimum number of edges that connect these large cities. The resulting route is depicted in Figure A7. As shown in the figure, the Euclidean spanning tree network is less precise in predicting today's road networks than the first two IVs, yet covers several important MSAs that are not on either IV's routes.

Figure A5. 1974 highway plan

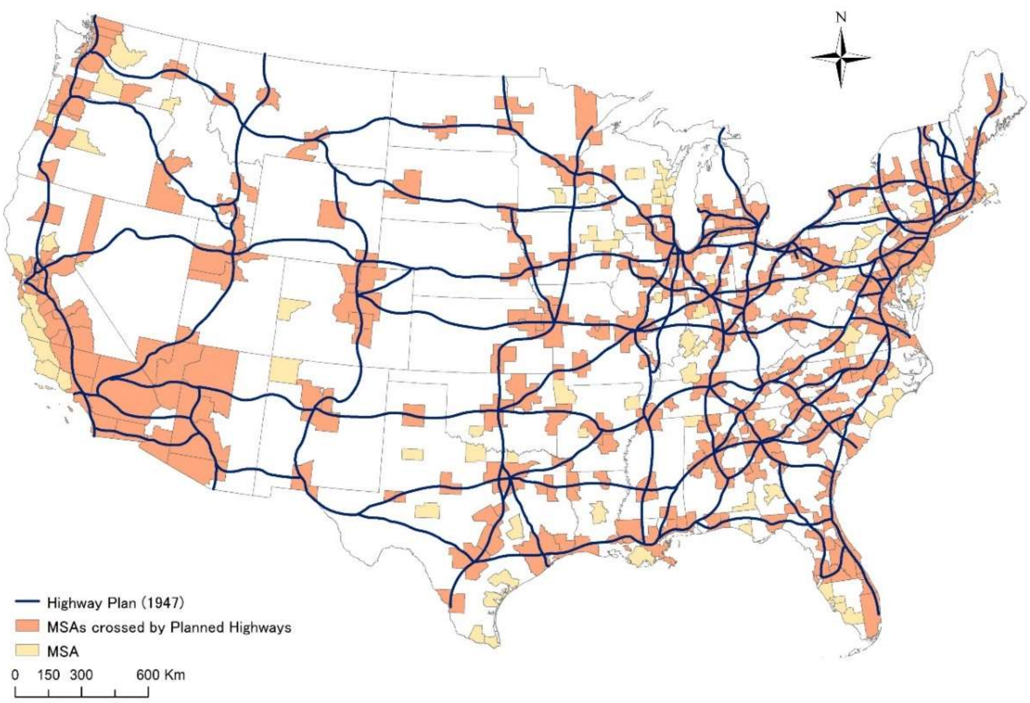


Figure A6. Railroad as of 1870

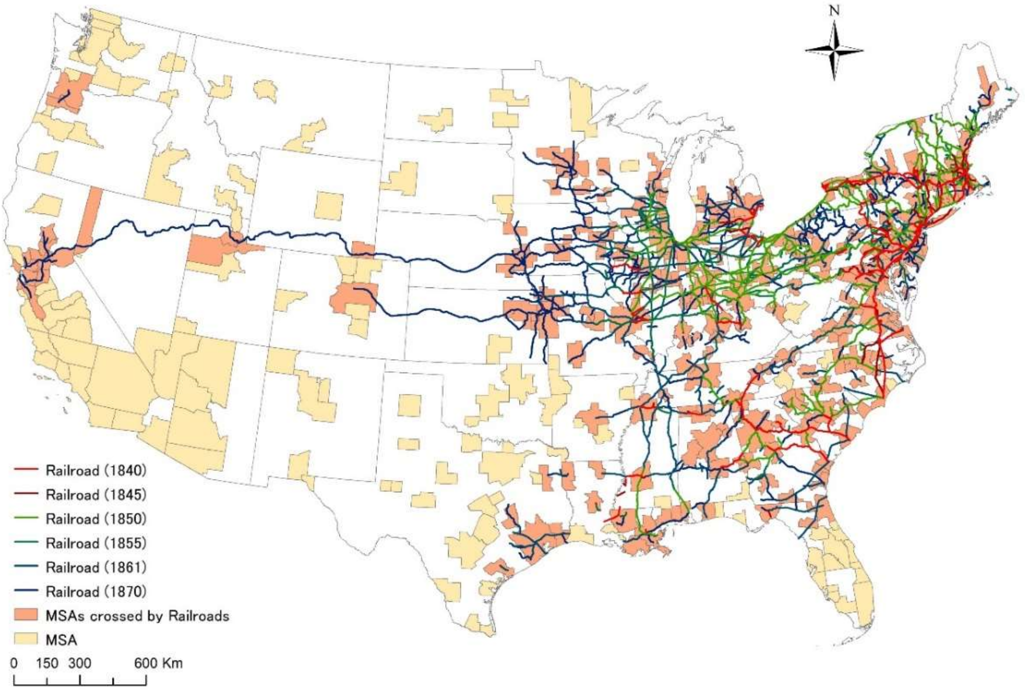
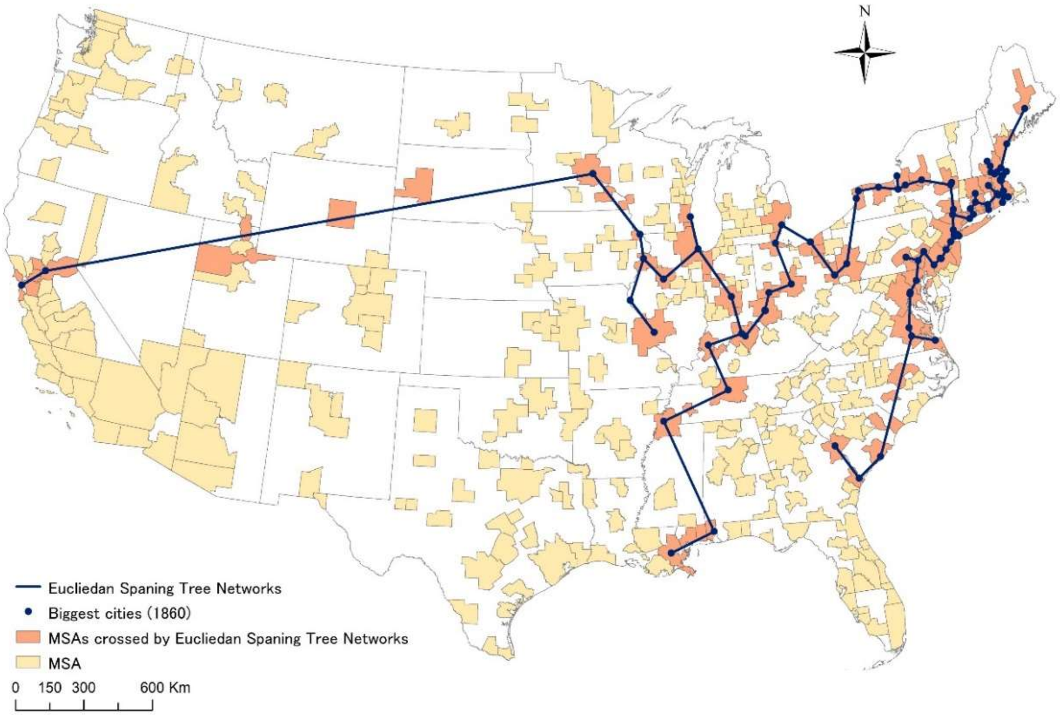


Figure A7. Euclidean network connecting large cities as of 1860



Appendix G. IV Falsification Test

We run the following regression to examine if our IVs are likely to satisfy the exclusion restriction.

$$Y_{ctm} = \alpha_c + \lambda_{tm} + \sum_k \beta^k \tilde{Z}^k_{ctm} + X'_{ctm}\gamma + \theta_c(t) + \epsilon_{ctm}$$

In this regression, our main outcome variables, the monthly means and maximums of ambient NO_2 concentration levels, are regressed on our three IVs on a subsample of MSAs with no ride-hailing entry, with the same set of full controls used in Table 2 in the main text. If the exclusion restriction holds, our IVs should not significantly affect the outcomes for these MSAs.

Note, however, that by definition, there are no reported entry dates for these MSAs. Hence, we create fictitious entry dates by using year dummies. Because most entry occurs in 2014, we use year 2013, 2014, and 2015 dummies and run each regression separately.

$$\tilde{Z}^k_{ctm} = G^k_c \times I_t(t = year)$$

The following table displays the results from this falsification test. We see that the F statistics on excluded instruments are statistically insignificant for all regressions. Hence, we conclude that our IVs pass the falsification test and, thus, are unlikely to be correlated with other unobserved confounds. There is a sign, however, that one of our IVs, the railroad-based instrument, is weakly correlated with the monthly maximums and in the non-urban areas. We, thus, interpret our IV results cautiously, particularly when using the monthly maximums as the main outcome variable.

Table A2. Falsification Test (Reduced-form Regressions on Non-entry MSAs)

	Mean NO_2 Concentration			**Max NO_2 Concentration**		
	Urban	**Suburban**	**Non-urban**	**Urban**	**Suburban**	**Non-urban**
***F* Statistics on Excluded Instruments**						
Year 2013	0.33	0.51	0.22	2.42	1.74	2.73
	(0.804)	(0.687)	(0.883)	(0.141)	(0.236)	(0.114)
Year 2014	0.65	1.08	2.14	2.38	1.73	2.61
	(0.603)	(0.410)	(0.174)	(0.146)	(0.238)	(0.123)
Year 2015	0.56	0.84	0.69	1.70	1.99	1.01
	(0.655)	(0.511)	(0.581)	(0.244)	(0.194)	(0.439)
Obs.	5,791	5,788	5,776	5,791	5,788	5,776
Adj. R^2	0.742	0.749	0.778	0.597	0.618	0.69

Note: P-values in parenthesis.