Is Ride-sharing Good for Environment?

Yoshifumi Konishi^{*} Department of Economics Keio University

and

Akari Ono Department of Economics Keio University

Draft: July 2022

^{*}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* and 21H00712). We thank Sho Kuroda, Takuya Shimamura, and Hazuki Yanagida for their valuable assistance. 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 NO₂ concentration data in the staggered difference-in-differences research design. Our empirical strategy accounts for treatment effect heterogeneity both within and across cities while controlling for pollution spillovers and exogenous time-varying covariates. We also employ 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 lower density cities. In the supplementary analysis using household-level commuting mode data, we 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) and from reducing information asymmetry (Liu *et al.*, 2018) and drivers' gain from Uber's compensation scheme (Angrist *et al.*, forthcoming) 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), 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).

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 (DID) design, exploiting the variation in ride-hailing entry over time and across cities. Our basic empirical strategy is to estimate this TWFE-DID 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.*'s (2018), Li *et al.* (2017), Ward *et al.* (2019), Sarmiento and Kim (2022)].

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 (NO₂) 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 NO₂ concentration data to calculate the ambient air pollution on spatially delineated subareas within each MSA. This allows us to compare, for example, urban-area NO₂ 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 DID 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 DID (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 also employ the Callaway-Sant'Anna estimator (2021) to estimate heterogeneous dynamic treatment effects. 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 estimate the TWFE-DID regression only on Texas MSAs, using this incidence as a quasi-experiment. Our last strategy applies an instrumental variable (IV) method to the TWFE-DID 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 NO₂ concentrations (in terms of both monthly means and maximums) for high population density MSAs, particularly in their urban areas. We also find no evidence of ride-hailing entry leading to an increase in ambient NO₂ concentrations in lower density MSAs. We also confirm these findings in the event study of dynamic treatment effects using a version of the Callaway-Sant'Anna 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 quintile are always negative, have relatively small standard errors, and range from -0.037 log points (TWFE-DID) to -0.061 log points (DD-IV). Assuming a linear relationship between satellite-based and EPA monitoring records, these impacts translate into the reductions of monthly mean NO₂ concentrations by roughly 0.57-0.95 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 NO₂ pollution. This reduction, however, is caused by a further decrease in private car commuters and increases in public transit commuters and in the 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, and Li *et al.* (2017), who show that Uber's entry decreases congestion.

Our work complements several vibrant areas of research: (a) empirical studies that estimate the causal effects of ride hailing on various economic outcomes [Angrist *et al.* (forthcoming); Berger *et al.* (2018); Chen *et al.* (2019); Cohen *et al.* (2018); Hall *et al.* (2018); Liu *et al.* (2018)], (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.* (Forthcoming); 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 [Fréchette *et al.* (2019); Hall *et al.* (2020); Shapiro (2020)].

Of these, our work is most closely related to Sarmiento and Kim (2022), who also use the Callaway-Sant'Anna estimator in a similar empirical design to ours and find that Uber's entry is estimated to improve air quality (as measured in the Air Quality Index and groundlevel ozone), particularly during the summer when bad air quality episodes are expected. Our study complements 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 difference-in-differences 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. We, thus, believe that our study reinforces their findings.

Lastly, the most important contribution of ours is that 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 that make a counterfactual comparison between hired rides versus other transit modes that could replace the rides tend to find that ride-hailing decreases use of mass transit and increases congestion [Erhardt et al. (2019), Kong et al. (2020), Rayle et al. (2016)]. On the other hand, studies that are based on quasi-experimental variations in ride hailing entry across MSAs or states tend to find that ride-hailing increases use of mass transit (Hall et al., 2018), decreases congestion (Li et al., 2017), and decrease ambient air pollution (Sarmiento and Kim, 2022) and vehicle emissions (Ward *et al.*, 2019). Our results suggest that ride-hailing can 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). We emphasize, however, that our results are also in line with the general recommendations the critics stress: To maximize ride-hailing's pollution reduction potential, local and federal governments need to increase investment in mass transit, encourage car pooling, and increase use of low- or zero-emissions vehicles by ride providers. Indeed, our results — both *entry* and *reduced supply* of ride-hailing service tends to improve ambient air quality in highly dense areas — suggests that there may be an optimal level of supply of ride-hailing service (or alternatively, an optimal level of congestion/Pigouvian pricing) for a given level of public transit network. 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 NO₂. 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 ride-hailing is alleged for 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)].

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.* (2017); Ward *et al.* (2019)] tend to refute the critics' arguments.¹

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 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?

Residents in a city consist of both car owners and non-car owners. When ride-hailing service enters the city, some of the car owners may subscribe to the ride-haling platform and become either net suppliers or net demanders of the service. Because any car owners can become occasional ride providers, the net expected return to car ownership would also increase, which may induce non-car owners to be car owners. A general consensus in the literature, both empirically and theoretically, is that the new ride-hailing service platform induces both a higher demand for ride service and a larger mass of service drivers, leading to an overall increase in the supply of ride-hailing driving hours (which includes deadheading).

¹An 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.

In fact, it is relatively straightforward to pin down conditions under which ride-hailing entry increases hours of driving by ride-hailing/taxi drivers in equilibrium, following the work of Hall *et al.* (2020).² Thus, total hours of driving (and transport-related emissions) in the city would unambiguously increase unless there is substantial substitution away from private driving.

Given this, some backgrounds on U.S. commuting patters are worth reviewing. 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. These examples suggest 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 illustrate these points in **Figure 1**. The top panel shows major train and bus routes in the urban and suburban areas of Chicago. In the bottom panel, we create a stylized diagram, which is intended to conceptualize the part of the city we highlight in the top

²Although quite unrealistic in the U.S. context, it indeed happened in some other parts of the world that ride-hailing entry had no effect on total supply. For example, in Japan, the local taxicab industry successfully lobbied against the entry of Uber into virtually all local markets in Japan, except in a few local municipalities. In these limited municipalities, the fare was kept extremely low to serve only the captive riders who cannot drive themselves. Such an extremely low fare virtually killed the ride-hailing market in Japan.

panel. Consider a commuter who lives in a suburb and commute to a city's center for work. There is a mass transit that can connect her to the station near her workplace. 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.

In sum, it is largely an empirical question whether or not such complementarity exists, and if exist, is strong enough to outweigh the increase in the supply of driving hours by rideservice 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. Our attempt to get at the effect of ride-hailing entry on transport-related emissions must account for its impact on such MSA-level commuting patterns.

3. Estimation and Identification Strategy

3.A. Overview

Our empirical strategy employs a staggered difference-in-differences (DID) research design, building on earlier empirical studies [Berger *et al.* (2018); Hall *et al.* (2018); Li *et al.* (2017)]. 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 DID 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:

$$y_{ctm}^{j} = \alpha_{c} + \lambda_{tm} + \sum_{s} \delta_{s} D_{ctm} + X_{ctm}^{\prime} \gamma + \theta_{c}(t) + \epsilon_{ctm}, \qquad (1)$$

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 quintile of MSA-level population density with $s \in \mathcal{S} \equiv \{1, 2, ..., 5\}$, 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 NO₂ 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 reported dates may have incomplete coverage, potentially multiple dates of entry, and some

 $^{^{3}}$ 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.

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 NO₂ concentration data, which are available at relatively high resolutions ($0.25^{\circ} \times 0.25^{\circ}$ girded cells) consistently throughout our study period. We discuss our NO₂ data more thoroughly in **Section 4**.

Third, although NO₂ 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 NO₂ concentration statistics for each of these subareas $j \in \{\text{urban}, \text{suburban}, \text{non-urban}\}$, and run a separate regression for each j in eq. (1). This allows us to compare, for example, urban-area NO₂ concentration of entry city against urban-area NO₂ 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 treatment effects differ over subareas of cities is also an important empirical question.

⁴It is known that the effect of emissions from on-road vehicles on ambient NO₂ 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 NO₂ 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.

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 DID 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 $\Omega_{ct} \ (\equiv \alpha_c, \lambda_{tm}, X_{ctm}, \theta_c(t))$, parallel trends in unobservables hold for all MSAs in the same population density cohort.

$$E[\epsilon_{c,t} - \epsilon_{c,t-1} | D_{ct} = 1, \Omega_{ct}] = E[\epsilon_{c,t} - \epsilon_{c,t-1} | D_{ct} = 0, \Omega_{ct}], \quad \forall \ c \in s, \ \forall \ s \in \mathcal{S}$$

Note that in the staggered DID 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**). 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). Sloczyński (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 DID setup, particularly when the effects are both dynamic and heterogeneous (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, Callaway-Sant'Anna (CS) estimator is 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 idea proposed by Gardner (2021) that one 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:

$$y_{ctm}^{j} = \alpha_{c} + \lambda_{tm} + X_{ctm}^{\prime}\gamma + \theta_{c}(t) + \epsilon_{ctm}$$

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

$$\hat{\epsilon}_{ctm}^{j} = y_{ctm}^{j} - (\hat{\alpha}_{c} + \hat{\lambda}_{tm} + X_{ctm}' \hat{\gamma} + \hat{\theta}_{c}(t))$$

3. Apply the following 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 t-th period as⁵

$$ATT^{j,s}(g,t) = E\left[\left(\frac{G_g}{E[G_g]} - \frac{C}{E[C]}\right)\left(\hat{\epsilon}_t^j - \hat{\epsilon}_{g-1}^j\right)\right]$$
(2)

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 t-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, 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 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 they are correlated with treatment assignment. We remove the bias by estimating the parameters using pre-treatment observations (Gardner, 2021). 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

⁵This is a version of Callaway-Sant'Anna estimator without covariate adjustment. 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 CS estimator separately on each density-quintile subsample.

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 ridehailing 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 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

⁶We, 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.

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 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,

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

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 quasiexperimental 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 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

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

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 *reported* entry dates of UberX, taken from Hall *et al.* (2018). We see that reported entry precedes *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. Thus, readers should be cautious in comparing our study with others because we differ from Berger *et al.* (2018), Hall *et al.* (2018) or Li *et al.* (2017) in terms of the measure of ride-hailing entry (so is the 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 two entry measures.⁸ There are several important take-away messages from the figure. First, the estimated impacts of reported entry are 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 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 NO₂ in molecules per cm² 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

⁸We use the Callaway-Sant'Anna estimator (2021) without covariates.

of other confounders on air pollution. NO₂ is a leading cause of respiratory diseases such as asthma, is a known precursor to ozone (O₃), and is often used as the indicator for the larger group of nitrogen oxides (NO_x). Because NO₂ is released into the air from combustion of fossil fuels, the primary sources of NO₂ 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 NO₂ emissions, but accounts for only 2-4% of PM_{2.5}, PM₁₀, and VOC emissions in the United States. The share of road transportation gets even higher in urbanized areas. For example, road transport-related air pollution, we should expect to see it most vividly in NO₂ concentration levels.⁹

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 Censustract-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

⁹Previous 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).

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 the **appendix Figure A2**, which plots the satellite-based NO₂ 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 NO₂ 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 NO₂, major changes to the NAAQS occurred in 1997, 2006, and 2012 for PM2.5 and in 1997, 2008, and 2015 for O₃. Because NO₂ is a known precursor to both pollutants, the nonattainment status for either pollutant may affect the time trends for NO₂ concentrations at the MSA level. Thus, we control for these in our regression.¹⁰

¹⁰There 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 NO₂ directly

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 NO₂ 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 entry here 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 NO₂ 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).¹¹ 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 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 O₃ and PM_{2.5}. Inclusion of endogenous time-varying controls can potentially bias our estimates since they can be correlated with

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 NO₂ may also affect ambient levels of PM2.5 or O₃, 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 NO₂ are smaller when we include these controls.

¹¹We use Sergio Correia's regulate package in Stata to efficiently absorb two-way fixed effects.

unobservables in the estimating equation. Hence, we avoid use of such controls. Standard errors are clustered at the MSA/year level.

In both tables, there is a tendency that the estimated impacts get smaller in magnitude for all regressions, suggesting that our estimates may be biased away from zero. 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 NO₂ for cities in the highest population density quintile. The estimates are statistically significant, and range from -0.034 log points (in non-urban areas) to -0.037 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 NO₂ concentrations by 0.53-0.58 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 quintile. 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.

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 quintile; the estimate get smaller for cities in the lower density cohorts, and turn positive for cities in the lowest density quintile. There is one important difference, however. The estimated impacts are larger, more statistically significant, and range from -0.053 log points (in non-urban areas) to -0.055 log points (in urban areas) in Table 2-B. 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 Sarmiento and Kim (2022), who finds that much of the air quality 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 Callaway-Sant'Anna estimator. As explained in **Section 3**, we implement the CS estimator separately for each density quintile (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 'control' 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 ATT (average treatment effect on the treated) for each quarter in relative time to entry (t = 0). To avoid busy graphs, we only present the results for the highest (5th) and the medium (3rd) population density quintiles. 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. The figure uses the monthly maximums as the outcome. The results using the monthly means are available in the **Appendix**.

In line with the results from the TWFE regression, we see that NO₂ 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 NO₂ concentrations (relative to not-yet-treated units) for the medium density cohort. The graph also indicates 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 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 householdlevel 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.¹² 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_o is a good proxy for the use of ride-sharing, $I_o \approx I_{rs}$, 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 ridehailing 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 commuting mode (e.g., car-pooling) 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 $(I_o \uparrow)$

¹²There 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 did not change much.

while commuting by private car stays roughly constant ($\Delta I_c \approx 0$). This is the case of ridehailing 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 $(I_c, I_{pt}, \text{ and } I_o)$. In all regressions, we include the MSA and year fixed effects as well as householdlevel 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 $(I_{walk}, I_{home}, \text{ and } I_{other})$. In the Appendix, we also report the results on two other outcomes (the number of passengers per ride and the time it takes to get to work) and the results are consistent with the results reported here. 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 0.6 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 fourth quintile 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 third quintile MSAs, we essentially have no impact of ride-hailing on commuting modes. For the last two quintiles, we see that ride-hailing entry is associated with increases in private car commuting, although the estimates are not statistically significant. 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 to see 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 implies 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. This is indeed consistent In the **Appendix**, we also show that the ride-hailing entry is also associated with increases in both the number of passengers per ride and the time it takes to get to work in the highest density MSAs.

Interestingly, in the fourth quintile MSAs, 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 (+0.6 ppt) is roughly the same as that of the reduction in the share of commuting by private car (-0.5 ppt). On one hand, these two effects tend to cancel out each other in terms of their impact on mean NO₂ 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 NO₂ 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 Sarmiento and Kim (2022).

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-experiments 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. This caused a sharp, unanticipated drop in the supply of ride-hailing services. 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. Importantly, this incidence occurred only in Austin, but no other neighboring MSAs. Figure 6-A confirms this point. There is one caution here — there is still a sign of some ride-hailing activity in Austin during this period. There is a small service called GetMe in Austin, who continued to provide the ride-hailing service during this period, and therefore, Uber and Lyft customers may have switched to their service. Hence, we take the FR as an event that caused a sharp decline 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. **Figure 6-B** helps us visualize our point. The figure plots the unconditional LOWESS regression of mean NO₂ concentrations on time periods, separately before and after Austin's ride-hailing entry date for Austin and no-entry MSAs. The figure indicates that mean NO₂ concentrations have an upward trend before the ride-hailing entry but switch to a declining trend after the entry, whereas no visible change is observed for no-entry MSAs. 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 A3**). Therefore, focusing on Texas allows us to control for state-level confounds while exploiting variation in timing and location of ride-hailing entry.

One disadvantage, however, is that the small sample size precludes the use of certain estimation strategies for reliable statistical inference. In particular, we cannot reliably use Callaway-Sant'Anna estimator because there are not enough 'relevant' MSAs for each treatment timing for each density cohort. We, therefore, estimate the TWFE regression in (1) on Texas MSAs while adding another FR dummy, which equals 1 for Austin during the FR period.

Table 4 presents the results on NO_2 concentrations. The results are very similar to those of the full sample reported in **Table 2**. The estimates get smaller with more controls and in less urban subareas (within the same density cohort). The estimated impacts of ride-hailing entry are negative and larger for MSAs with higher population density. In line with our expectation, the estimates have larger standard errors in Table 4 than Table 2, and thus, are less statistically significant. Interestingly, the fingerprint rule is estimated to decrease NO_2 concentrations, and the estimated impacts are large (even relative to the large standard errors). Note, however, 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 quintile, we obtain the pure impacts of the FR by subtracting the estimates for that quintile: 0.050-0.055 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 auxiliary TWFE regressions using yearly householdlevel 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 no change 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 unclear, 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 DID 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_{ct} 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 geographybased 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**.

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_{ct}^k for city c and time t is:

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

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 ($D_{ct} \geq R_{ct}$). 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 pretreatment covariates such as the current population size. Identification is strong because we interact reported entry with geography-based instruments, and thus, we only use the timing of reported entry, not entry location, as information for constructing our instruments.

For implementation, we apply the two-stage least squares (2SLS) estimator on the DID 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: $E[D_{ct}|Z_{ct}=1] \ge E[D_{ct}|Z_{ct}=0]$.
- 2. Exclusion restriction: Z_{ct} affects Y_{ct} only through D_{ct} .
- 3. Conditional parallel trends in Y_{ct} on Z_{ct} : Conditional on exogenous (time-varying) covariates Ω_{ct} ($\equiv \alpha_c, \lambda_{tm}, X_{ctm}, \theta_c(t)$), parallel trends in unobservables hold for all MSAs in the same population density cohort s.

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

Table 6 reports the DD-IV estimation results on mean NO₂ concentrations, using all three geography-based instruments interacted with the reported ride-hailing entry dummy. We use the same set of controls and 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 because the 2SLS estimation generally leads to larger standard errors and 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 5th and 4th quintiles), particularly in urban areas. The signs of the estimates turn positive for medium- and lowest-density MSAs. The estimated effects of ride-hailing entry on the 2nd quintile MSAs are negative, but this is still consistent with the TWFE regression. We take these as supportive evidence for our main results.

The magnitudes of the estimates seem economically too large, however, particularly for lower density MSAs. For example, ride-hailing entry is estimated to increase mean urbanarea NO_2 concentrations by 0.213 log points in the lowest quintile and to decrease by 0.168 log points in the 2nd quintile. 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, 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 fail to 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 not be biased as long as the other identifying assumptions hold.

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**. Interestingly, the estimated impacts

on commuting modes slightly differ from the TWFE estimates in **Table 3**, yet are indeed consistent with the estimated impacts on NO₂ concentrations in **Table 6**. For the highest density MSAs, ride-hailing entry is estimated to decrease private car commuters by 4.3 ppt, but much of this decline is associated with an increase in public transit commuters (by 3.5 ppt) rather than in other modes. In contrast, for the 4th quintile MSAs, a decrease in private car commuters (by 3.4 ppt) is accompanied by an increase in the share of commuters using other modes (by 2.6 ppt), and much of this change comes from an increase in the share of working at home (by 2.2 ppt). Furthermore, ride-hailing entry is estimated to reduce public transit commuters by 0.5 ppt for the lowest and 2nd quintile MSAs, yet this decline is associated with a larger increase in private car commuters in the lowest quintile MSAs (by 0.9 ppt) and a similar increase in the share of those walking to work (by 0.5 ppt). Hence, taking these results at the face value, ride-hailing entry seems to affect commuting patterns quite differently across different MSAs, and these differences seem to explain well its differential air-pollution impacts across different MSAs.

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 alternative 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, which shows that private car ownership decline, 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.* (Forthcoming)]. In a study covering 58 subway openings worldwide, for example, Gendron-Carrier *et al.* (Forthcoming) 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 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

- 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
- [2] Arnott, Richard. (1996) Taxi Travel Should Be Subsidized. Journal of Urban Economics 40 (3): 316-333
- [3] Angrist, Joshua D., Sydnee Caldwell, Jonathan V. Hall. (2021) Uber vs. Taxi: A Driver's Eye View. Forthcoming at American Economic Journal: Applied Economics
- [4] 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
- [5] Basker, Emek (2005) Job Creation or Destruction? Labor Market Effects of Wal-Mart Expansion. The Review of Economics and Statistics 87 (1): 174–183
- [6] Baum-Snow, Nathaniel. (2007) Did Highways Cause Suburbanization? The Quarterly Journal of Economics 122 (2): 775–805

- [7] Berger, Thor, Chinchih Chen, Carl Benedikt Frey. (2018) Drivers of Disruption? Estimating the Uber Effect. European Economic Review 110: 197–210
- [8] Callaway, Brantly and Pedro H.C. Sant'Anna. (2020) Difference-in-Differences with Multiple Time Periods. *Journal of Econometrics* in press.
- [9] 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
- [10] 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
- [11] 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
- [12] 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
- [13] Correia, Sergio. (2017) Linear Models with High-Dimensional Fixed Effects: An Efficient and Feasible Estimator. Working Paper. http://scorreia.com/research/hdfe.pdf.
- [14] de Chaisemartin, Clement, Xavier D'Haultfoeuille (2020) Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects. American Economic Review
- [15] 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
- [16] Dills, Angela K., Sean E. Mulholland. (2018) Ride-Sharing, Fatal Crashes, and Crime. Southern Economic Journal 84 (4): 965–991
- [17] 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.
- [18] Duranton, Gilles, Matthew A. Turner. (2011) The Fundamental Law of Road Congestion: Evidence from US Cities. American Economic Review 101: 2616–2652
- [19] Duranton, Gilles, Matthew A. Turner. (2012) Urban Growth and Transportation. The Review of Economic Studies 79 (4): 1407–1440

- [20] Erhardt, Gregory D., Sneha Roy, Drew Cooper, Bhargava Sana, Mei Chen, Joe Castiglione. Do Transportation Network Companies Decrease or Increase Congestion? *Sci*ence Advances 5 (5), aau2670 DOI: 10.1126/sciadv.aau2670
- [21] 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
- [22] 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
- [23] 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-pollutionin-europereport
- [24] 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
- [25] Gardner, John (2021) Two-stage Differences in Differences. Working paper
- [26] Gendron-Carrier, Nicolas, Marco Gonzalez-Navarro, Stefano Polloni, Matthew A. Turner (2021) Subways and Urban Air Pollution. Forthcoming at American Economic Journal: Applied Economics
- [27] Goodman-Bacon, Andrew. (2021) Difference-in-Differences with Variation in Treatment Timing. Journal of Econometrics in press.
- [28] Grainger, Corbett, Andrew Schreiber (2019) Discrimination in Ambient Air Pollution Monitoring? AEA Papers and Proceedings 109: 277-82.
- [29] Gu, Yizhen, Chang Jiang, Junfu Zhang, and Ben Zou (2021) Subways and Road Congestion. American Economic Journal: Applied Economics 13 (2): 83–115
- [30] Hall, Jonathan D., Craig Palsson, Joseph Price (2018) Is Uber A Substitute or Complement for Public Transit? *Journal of Urban Economics* 108: 36-50
- [31] Hall, Jonathan V., John J. Horton, Daniel T. Knoepfle (2020) Ride-Sharing Markets Re-Equilibrate. Working Paper

- [32] 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
- [33] 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
- [34] Lagos, R. (2003) An Analysis of the Market for Taxicab Rides in New York City. International Economic Review 44 (2): 423-434
- [35] 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,
- [36] Li, Ziru, Yili Hong, Zhongju Zhang. (2016) Do On-demand Ride-sharing Services Affect Traffic Congestion? Evidence from Uber Entry. Available at SSRN: https://ssrn.com/abstract=2838043
- [37] Liu, Meng, Erik Brynjolfsson, and Jason Dowlatabadi. (2021) Do Digital Platforms Reduce Moral Hazard? The Case of Uber and Taxis. NBER Working Paper 25015
- [38] 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
- [39] 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
- [40] Neumark, David, Junfu Zhang, Stephen Ciccarella (2008) The Effects of Wal-Mart on Local Labor Markets. *Journal of Urban Economics* 63 (2): 405-430
- [41] 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
- [42] Słoczyński, Tymon (2020) Interpreting OLS Estimands When Treatment Effects Are Heterogeneous: Smaller Groups Get Larger Weights. *Review of Economics & Statistics*

- [43] 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
- [44] Sun, Liyang, Sarah Abraham (2020) Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects. *Journal of Econometrics*
- [45] 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
- [46] 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
- [47] Research Triangle Park, North Carolina. # EPA-452/R-08-008a
- [48] 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
- [49] Wooldridge, Jeffrey M. (2021) Two-Way Fixed Effects, the Two-Way Mundlak Regression, and Event Study Estimators. Working paper
- [50] 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





(a) San Francisco, CA









(a) Share of MSAs with Uber/Lyft Entry





B. Austin and San Antonio, TX



C. San Francisco, CA







Figure 6. Quasi-experiments in Texas



(a) Normalized Google Trends for Texas MSAs





	No Ei	ıtry	Entry ir	1 2013	Entry i	n 2014	After	2015	Α	1
Ι	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
N. of MSAs	62		3((21	7	36	(34	8
(Share of full sample)	(0.1)	8)	(0.0)	(6)	(0.0	52)	(0.1	(1)	(1.0	(0)
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	0.0	×	0.2	L.	0.0	22	0.0	15	0.1	×

S
5
, S
5
÷
n
le
E
It
Ğ
Ē
► N
S
÷
ij
P
Ξ
õ
\mathbf{U}
ut i
- E
B
5
8
Ľ
÷
ف
2
-
دە
Ť
- P
<u></u>

Note: All statistics are as of year 2010.

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

				Mean	NO2Concent	ratio n			
		Urban Area		s	uburban Are	a	N	on-urban Are	a
Entry									
× 1st Quintile MSAs	0.009 (0.017)	0.014 (0.018)	0.014 (0.018)	0.009 (0.016)	0.014 (0.017)	0.014 (0.017)	0.029 (0.018)	0.036 * (0.018)	0.037 * (0.019)
\times 2nd Quintile MSAs	-0.016 (0.012)	-0.009 (0.011)	-0.008 (0.012)	-0.015 (0.011)	-0.008 (0.011)	-0.007 (0.012)	-0.005 (0.011)	0.003 (0.011)	0.005 (0.013)
\times 3rd Quintile MSAs	-0.016 (0.015)	-0.011 (0.015)	-0.009 (0.015)	-0.014 (0.013)	-0.010 (0.014)	-0.007 (0.013)	-0.013 (0.013)	-0.008 (0.014)	-0.003 (0.013)
\times 4th Quintile MSAs	-0.019 (0.014)	-0.015 (0.015)	-0.012 (0.014)	-0.021 (0.013)	-0.016 (0.014)	-0.013 (0.013)	-0.020 (0.012)	-0.016 (0.013)	-0.012 (0.012)
\times 5th Quintile MSAs	-0.048 ** (0.014)	-0.040 ** (0.014)	-0.037 ** (0.013)	-0.046 ** (0.014)	-0.039 ** (0.014)	-0.036 ** (0.013)	-0.045 *** (0.013)	-0.039 ** (0.013)	-0.034 ** (0.013)
Controls included:									
Year/ Month FE	\checkmark	\checkmark	\checkmark						
MSA FE	\checkmark	\checkmark	\checkmark						
Climate controls		\checkmark	\checkmark		\checkmark	\checkmark		\checkmark	\checkmark
Post-tr. controls			\checkmark			\checkmark			\checkmark
Obs. Adj. R ²	38,911 0.759	35,023 0.757	35,023 0.757	39,309 0.765	35,421 0.762	35,421 0.762	39,254 0.781	35,367 0.778	35,367 0.778

A. Monthly Mean NO₂ Concentrations

Note: Clustered robust standard errors in parenthesis.

B. Monthly Maximum NO₂ Concentrations

				Max	NO2 Concent	ration			
		Urban Area		5	uburban Are	a	N	on-urban Are	ea
Entry									
× 1st Quintile MSAs	0.001 (0.020)	0.002 (0.021)	0.003 (0.021)	0.000 (0.018)	0.002 (0.019)	0.002 (0.019)	0.019 (0.016)	0.023 (0.017)	0.024 (0.017)
\times 2nd Quintile MSAs	-0.027 (0.016)	-0.023 (0.016)	-0.022 (0.016)	-0.025 (0.016)	-0.021 (0.016)	-0.020 (0.017)	-0.016 (0.012)	-0.012 (0.012)	-0.009 (0.014)
\times 3rd Quintile MSAs	-0.028 (0.016)	-0.026 (0.017)	-0.023 (0.017)	-0.029 (0.016)	-0.027 (0.016)	-0.024 (0.016)	-0.028 * (0.012)	-0.026 * (0.013)	-0.021 (0.013)
\times 4th Quintile MSAs	-0.035 ** (0.015)	-0.029 * (0.015)	-0.025 (0.014)	-0.037 ** (0.015)	-0.032 * (0.015)	-0.027 * (0.013)	-0.038 ** (0.013)	-0.033 ** (0.013)	-0.028 ** (0.012)
× 5th Quintile MSAs	-0.065 ** (0.023)	-0.060 ** (0.022)	-0.055 ** (0.020)	-0.064 ** (0.023)	-0.058 ** (0.021)	-0.054 ** (0.020)	-0.064 ** (0.021)	-0.059 ** (0.019)	-0.053 ** (0.019)
Controls included:									
Year/ Month FE	\checkmark								
MSA FE	\checkmark								
Climate controls		\checkmark	\checkmark		\checkmark	\checkmark		\checkmark	\checkmark
Post-tr. controls			\checkmark			\checkmark			\checkmark
Obs. Adj. R ²	38,911 0.662	35,023 0.66	35,023 0.661	39,309 0.679	35,421 0.677	35,421 0.677	39,254 0.712	35,367 0.709	35,367 0.709

Note: Clustered robust standard errors in parenthesis.

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

				Com	muting Mode t	o Work			
		Private Car			Public Transi	it		Other Mode	
Entry									
× 1st Quintile MSAs	0.001 (0.002)	0.003 (0.002)	0.003 (0.002)	0.000 (0.001)	-0.000 (0.001)	-0.001 (0.001)	-0.001 (0.002)	-0.001 (0.002)	-0.001 (0.002)
× 2nd Quintile MSAs	0.000 (0.002)	0.002 (0.002)	0.002 (0.002)	-0.000 (0.001)	-0.000 (0.001)	-0.000 (0.001)	-0.000 (0.002)	-0.002 (0.002)	-0.002 (0.002)
× 3rd Quintile MSAs	-0.002 (0.002)	-0.000 (0.002)	0.000 (0.002)	-0.001 (0.001)	-0.001 (0.001)	-0.001 (0.001)	0.003 ** (0.001)	0.001 (0.001)	0.001 (0.001)
\times 4th Quintile MSAs	-0.007 *** (0.002)	-0.006 *** (0.002)	-0.005 **** (0.002)	-0.000 (0.001)	-0.001 (0.001)	-0.001 (0.001)	0.006 *** (0.001)	0.006 *** (0.001)	0.006 *** (0.001)
\times 5th Quintile MSAs	-0.015 *** (0.003)	-0.014 *** (0.003)	-0.013 **** (0.002)	0.005 * (0.003)	0.005 * (0.003)	0.004 * (0.002)	0.008 *** (0.001)	0.007 *** (0.001)	0.006 *** (0.001)
Controls included:									
Year FE	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark
MSA FE	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark
Demog. Controls	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark
Climate controls		\checkmark	\checkmark		\checkmark	\checkmark		\checkmark	\checkmark
Other controls			\checkmark			\checkmark			\checkmark
Obs. Adj. R ²	18,936,750 0.061	17,872,388 0.063	17,872,388 0.063	18,936,750 0.115	17,872,388 0.116	17,872,388 0.116	18,936,750 0.007	17,872,388 0.007	17,872,388 0.007

A. Primary Commuting Mode

Note: Clustered robust standard errors in parenthesis. Repeated cross-section, household-level commuting mode data from American Community Survey (ACS). Controls include household-level covariates.

B. Other Commuting Mode: Subcategories

				Other Co	mmuting Mod	le to Work			
		Walking			Work at Hom	e		Other Mode	2
Entry									
× 1st Quintile MSAs	-0.000 (0.001)	-0.000 (0.001)	-0.000 (0.001)	-0.000 (0.001)	-0.001 (0.001)	-0.001 (0.001)	0.000 (0.001)	0.000 (0.001)	-0.000 (0.001)
× 2nd Quintile MSAs	-0.001 (0.001)	-0.001 ** (0.001)	-0.001 ** (0.001)	0.000 (0.001)	-0.001 (0.001)	-0.001 (0.001)	0.000 (0.001)	0.000 (0.001)	-0.000 (0.001)
× 3rd Quintile MSAs	-0.000 (0.001)	-0.001 (0.001)	-0.001 (0.001)	0.004 *** (0.001)	0.002 ** (0.001)	0.002 ** (0.001)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
× 4th Quintile MSAs	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.006 *** (0.001)	0.006 *** (0.001)	0.006 *** (0.001)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
× 5th Quintile MSAs	0.002 *** (0.000)	0.002 *** (0.000)	0.001 *** (0.000)	0.006 *** (0.001)	0.005 *** (0.001)	0.005 *** (0.001)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
Controls included:									
Year FE	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark
MSA FE	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark
Demog. Controls	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark
Climate controls		\checkmark	\checkmark		\checkmark	\checkmark		\checkmark	\checkmark
Other controls			\checkmark			\checkmark			\checkmark
Obs. Adj. R ²	18,936,750 0.007	17,872,388 0.007	17,872,388 0.007	18,936,750 0.010	17,872,388 0.010	17,872,388 0.010	18,936,750 0.002	17,872,388 0.002	17,872,388 0.002

Note: Clustered robust standard errors in parenthesis. Repeated cross-section, household-level commuting mode data from American Community Survey (ACS). Controls include household-level covariates.

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

				Mean	NO2Concen	tration			
		Urban Area		s	uburban Are	a	N	on-urban Ar	ea
Entry									
× 1st Quintile MSAs	-0.008 (0.019)	-0.007 (0.042)	-0.014 (0.028)	-0.002 (0.018)	0.000 (0.067)	-0.006 (0.027)	-0.005 (0.017)	-0.004 (0.019)	-0.007 (0.021)
× 2nd Quintile MSAs	0.016 (0.018)	0.021 (0.017)	0.017 (0.019)	0.015 (0.016)	0.021 (0.015)	0.018 (0.017)	0.018 (0.015)	0.022 (0.015)	0.020 (0.016)
× 3rd Quintile MSAs	0.013 (0.012)	0.019 (0.017)	0.012 (0.017)	0.017 (0.012)	0.024 (0.014)	0.018 (0.017)	0.017 (0.015)	0.023 (0.019)	0.019 (0.019)
× 4th Quintile MSAs	-0.015 (0.016)	-0.011 (0.015)	-0.010 (0.019)	-0.005 (0.017)	-0.000 (0.015)	0.000 (0.019)	-0.004 (0.017)	-0.001 (0.015)	-0.003 (0.020)
× 5th Quintile MSAs	-0.029 (0.026)	-0.026 (0.029)	-0.018 (0.033)	-0.023 (0.026)	-0.020 (0.039)	-0.014 (0.033)	-0.008 (0.023)	-0.006 (0.022)	-0.009 (0.030)
Fingerprint Rule	-0.079 ** (0.029)	-0.077 (0.101)	-0.073 (0.074)	-0.069 ** (0.027)	-0.068 (0.129)	-0.065 (0.068)	-0.066 ** (0.025)	-0.065 (0.058)	-0.066 (0.061)
Controls included:									
Year FE	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark
Month FE	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark
MSA FE	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark
Climate controls		\checkmark	\checkmark		\checkmark	\checkmark		\checkmark	\checkmark
Other controls			\checkmark			\checkmark			\checkmark
Obs.	2,016	2,016	2,016	2,016	2,016	2,016	2,016	2,016	2,016
Adj. R^2	0.767	0.771	0.771	0.764	0.769	0.769	0.762	0.767	0.767

A. Monthly Mean NO₂ Concentrations

Note: Clustered robust standard errors in parenthesis.

B. Monthly Maximum NO₂ Concentrations

				Max N	O2 Concent	ration			
		Urban Area		s	uburban Are	a	Ν	on-urban Ar	ea
Entry									
× 1st Quintile MSAs	-0.008 (0.030)	-0.005 (0.030)	-0.010 (0.026)	-0.000 (0.020)	0.003 (0.027)	0.000 (0.022)	-0.001 (0.014)	0.003 (0.015)	0.000 (0.014)
× 2nd Quintile MSAs	0.011 (0.035)	0.017 (0.043)	0.014 (0.036)	0.005 (0.033)	0.012 (0.038)	0.010 (0.034)	0.007 (0.023)	0.012 (0.028)	0.011 (0.027)
× 3rd Quintile MSAs	-0.038 (0.037)	-0.031 (0.051)	-0.035 (0.056)	-0.026 (0.026)	-0.018 (0.064)	-0.020 (0.044)	-0.016 (0.022)	-0.009 (0.042)	-0.011 (0.049)
\times 4th Quintile MSAs	-0.054 *** (0.012)	-0.049 ** (0.013)	-0.043 ** (0.016)	-0.037 * (0.017)	-0.031 (0.023)	-0.026 (0.019)	-0.022 (0.016)	-0.017 (0.023)	-0.016 (0.022)
\times 5th Quintile MSAs	-0.079 * (0.033)	-0.075 ** (0.031)	-0.055 (0.036)	-0.075 ** (0.029)	-0.070 * (0.034)	-0.054 (0.037)	-0.043 * (0.020)	-0.039 (0.023)	-0.035 (0.031)
Fingerprint Rule	-0.123 ** (0.038)	-0.118 (0.086)	-0.124 (0.072)	-0.103 ** (0.036)	-0.099 (0.095)	-0.103 (0.070)	-0.098 ** (0.037)	-0.094 (0.071)	-0.095 (0.071)
Controls included:									
Year FE	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark
Month FE	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark
MSA FE	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark
Climate controls		\checkmark	\checkmark		\checkmark	\checkmark		\checkmark	\checkmark
Other controls			\checkmark			\checkmark			\checkmark
Obs. Adj. R ²	2,016 0.688	2,016 0.692	2,016 0.692	2,016 0.703	2,016 0.707	2,016 0.707	2,016 0.722	2,016 0.726	2,016 0.726

Note: Clustered robust standard errors in parenthesis.

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

				Comm	uting Mode to	Work			
		Private Car			Public Transit			Other Mode	
Entry									
× 1st Quintile MSAs	-0.002 (0.003)	-0.002 (0.003)	-0.002 (0.003)	0.001 ** (0.000)	0.001 ** (0.000)	0.002 *** (0.000)	0.002 (0.002)	0.002 (0.002)	0.001 (0.003)
× 2nd Quintile MSAs	0.003 (0.005)	0.003 (0.005)	0.003 (0.005)	-0.002 (0.003)	-0.002 (0.003)	-0.002 (0.003)	-0.002 (0.003)	-0.002 (0.003)	-0.003 (0.003)
\times 3rd Quintile MSAs	-0.015 *** (0.002)	-0.015 *** (0.002)	-0.015 *** (0.002)	-0.001 * (0.000)	-0.001 * (0.000)	-0.000 (0.000)	0.016 *** (0.002)	0.016 *** (0.002)	0.015 *** (0.002)
\times 4th Quintile MSAs	-0.002 (0.004)	-0.002 (0.004)	-0.002 (0.005)	-0.002 *** (0.000)	-0.002 *** (0.000)	-0.002 **** (0.000)	0.004 (0.004)	0.004 (0.004)	0.004 (0.004)
\times 5th Quintile MSAs	-0.006 ** (0.002)	-0.006 ** (0.002)	-0.006 ** (0.002)	-0.001 * (0.001)	-0.001 * (0.001)	-0.001 ** (0.001)	0.007 *** (0.002)	0.007 *** (0.002)	0.007 *** (0.002)
Fingerprint Rule	-0.010 **** (0.002)	-0.010 *** (0.002)	-0.011 **** (0.002)	0.004 *** (0.001)	0.004 *** (0.001)	0.005 **** (0.001)	0.006 *** (0.001)	0.006 *** (0.001)	0.006 *** (0.001)
Controls included:									
Year FE	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark
MSA FE	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark
Demog. Controls	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark
Climate controls		\checkmark	\checkmark		\checkmark	\checkmark		\checkmark	\checkmark
Other controls			\checkmark			\checkmark			\checkmark
Obs. Adj. R ²	968,841 0.007	968,841 0.007	968,841 0.007	968,841 0.007	968,841 0.007	968,841 0.007	968,841 0.007	968,841 0.007	968,841 0.007

A. Primary Commuting Mode

Note: Clustered robust standard errors in parenthesis.

B. Other Commuting Mode: Subcategories

				Other Co	mmuting Mode	to Work			
		Walking			Work at Home			Other Mode	
Entry									
× 1st Quintile MSAs	0.004 *** (0.002)	0.004 *** (0.002)	0.004 ** (0.002)	-0.003 (0.003)	-0.003 (0.003)	-0.003 (0.003)	0.000 (0.001)	0.000 (0.001)	0.000 (0.001)
\times 2nd Quintile MSAs	-0.001 (0.003)	-0.001 (0.003)	-0.001 (0.003)	0.001 (0.002)	0.001 (0.002)	0.000 (0.002)	-0.002 (0.002)	-0.002 (0.002)	-0.002 (0.002)
× 3rd Quintile MSAs	0.000 (0.001)	0.000 (0.001)	0.000 (0.001)	0.012 *** (0.002)	0.012 *** (0.002)	0.011 **** (0.002)	0.004 *** (0.001)	0.004 *** (0.001)	0.004 *** (0.001)
\times 4th Quintile MSAs	0.001 (0.001)	0.001 (0.001)	0.001 (0.001)	0.005 ** (0.002)	0.005 ** (0.002)	0.005 ** (0.002)	-0.002 (0.002)	-0.002 (0.002)	-0.002 (0.002)
× 5th Quintile MSAs	0.002 ** (0.001)	0.002 ** (0.001)	0.002 ** (0.001)	0.006 *** (0.002)	0.006 *** (0.002)	0.007 **** (0.002)	-0.002 (0.001)	-0.002 (0.001)	-0.002 (0.001)
Fingerprint Rule	-0.004 **** (0.000)	-0.004 *** (0.000)	-0.004 **** (0.000)	0.012 *** (0.001)	0.012 *** (0.001)	0.012 *** (0.001)	-0.002 **** (0.001)	-0.002 *** (0.001)	-0.002 *** (0.000)
Controls included:									
Year FE	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark
MSA FE	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark
Demog. Controls	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark
Climate controls		\checkmark	\checkmark		\checkmark	\checkmark		\checkmark	\checkmark
Other controls			\checkmark			\checkmark			\checkmark
Obs. Adi R ²	968,841	968,841	968,841	968,841	968,841 0.014	968,841	968,841	968,841	968,841

Note: Clustered robust standard errors in parenthesis.

		Po	pulation Densi	ty	
	Lowest	2nd	3rd	4th	Highest
	MSAs	MSAs	MSAs	MSAs	MSAs
Mean NO ₂					
Urban area	0.213	-0.168	0.053	-0.057	-0.061
	(0.190)	(0.134)	(0.060)	(0.081)	(0.074)
Suburban area	0.175	-0.170	0.057	-0.068	-0.005
	(0.189)	(0.134)	(0.059)	(0.080)	(0.072)
Non-urban area	0.101	-0.121	0.081	-0.116	0.019
	(0.178)	(0.122)	(0.061)	(0.079)	(0.070)
Obs.	6,649	7,078	6,847	7,157	7,028
Weak IV stat.	33.172	68.972	130.324	87.032	148.35
Hansen's J stat.	0.106	0.121	2.766	0.481	4.365
(p-values)	0.745	0.941	0.251	0.786	0.113

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

Note: Clustered robust standard errors in parenthesis. 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.

		Po	pulation Densi	ty	
	Lowest	2nd	3rd	4th	Highest
	MSAs	MSAs	MSAs	MSAs	MSAs
Commuting Mode to V	Work				
Private Car	0.009	-0.000	0.002	-0.034	-0.043
	(0.013)	(0.016)	(0.005)	(0.016)	(0.030)
Public Transit	-0.005	-0.005	0.000	0.003	0.035
	(0.003)	(0.006)	(0.002)	(0.004)	(0.028)
Other Mode	0.002	0.005	-0.003	0.026	0.007
	(0.011)	(0.012)	(0.005)	(0.013)	(0.012)
Obs.	1,141,576	1,587,313	1,798,198	3,287,572	9,967,071
Weak IV stat.	21,324	30,000	120,984	94,467	184,005
Hansen's J stat.	0.122	2.409	1.393	1.607	3.226
(p-values)	0.727	0.3	0.498	0.448	0.199

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

Note: Clustered robust standard errors in parenthesis. We use the full set of controls for all regressions. Weak identification/ overidentification statistics are reported for private car regressions only. Regressions are separately estimated for each population density strata.

Appendix A

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.

	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-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					

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

Appendix B

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 NO₂ concentrations in our main analysis.

Figure A1. Geographic Distribution of Ride-haling Entry and Urban Area Boundaries



Appendix C

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 NO₂ concentrations for each level of EPA NO₂ concentrations. There is also an indication that EPA data might be missing for high concentration episodes, as in Zou (2021) and Grainger (2021).





Appendix D

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





Appendix E. 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 A4** 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 A5** 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 A6**. 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 A4. 1974 highway plan



Figure A5. Railroad as of 1870

Figure A6. Euclidean network connecting large cities as of 1860



Appendix F. 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.

$$\tilde{Z}_{ctm}^k = G_c^k \times I_t(t = year)$$

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.

_	Mean NO2 Concentration			Max NO2 Concentration		
_	Urban	Suburban	Non-urban	Urban	Suburban	Non-urban
F Statistics on Excluded Instruments						
Year 2013	0.33 (0.804)	0.51 (0.687)	0.22 (0.883)	2.42 (0.141)	1.74 (0.236)	2.73 (0.114)
Year 2014	0.65 (0.603)	1.08 (0.410)	2.14 (0.174)	2.38 (0.146)	1.73 (0.238)	2.61 (0.123)
Year 2015	0.56 (0.655)	0.84 (0.511)	0.69 (0.581)	1.70 (0.244)	1.99 (0.194)	1.01 (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

Table A2. Falsification Test (Red	uced-form F	Regressions (on Non-entry	v MSAs)
-----------------------------------	-------------	---------------	--------------	---------

Note: P-values in parenthesis.