Is Ride-sharing Good for Environment?
Evidence from Combining Satellite and Survey Data on U.S. Cities

Yoshifumi Konishi*
Faculty of Economics
Keio University

and

Akari Ono
Faculty of Economics
Keio University

Draft: *May 2026*

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

Acknowledgement: The study was in part supported by the financial supports from Japan Society for the Promotion of Science (JSPS), Grant-in-aid for Scientific Research (*Grant numbers: 18K01562, 21H00712,* and 23K20599). Sho Kuroda, Takuya Shimamura, and Hazuki Yanagida provided valuable research assistance. We thank Michiko Namazu, a data scientist at Uber Inc., for sharing Uber's official launch dates in the U.S. We also thank Yuta Toyama, Toshi Arimura, and other seminar participants at Waseda University's Empirical Microeconomics Workshop for their comments. We are also grateful to three anonymous referees and the co-editor, Stephan Heblich, for their invaluable comments and suggestions on our manuscript. This is a substantially improved version of the April 2021 draft and the earlier discussion paper (Keio-IES DP2024-014).

Abstract: We estimate the causal effect of ride-hailing entry on transport-related air pollution, disentangling its mediating effect through changes in commuting modes in U.S. cities. To do so, we combine two sets of empirical approaches. First, for our main outcome regression, we leverage granular satellite-based NO_2 concentration data and a newly constructed Google Trends-based measure of ride-hailing presence, with the staggered difference-in-differences design. Second, to explore its mediating mechanism, we use household-level commuting mode data to run two auxiliary regressions: commuting modes on ride-hailing entry and ambient NO_2 concentration on commuting modes. For identification on the latter, we construct our instruments by combining geography-based instruments with leave-one-out regional average exposure to Uber's official entry. We find robust evidence that (i) ride-hailing improves air quality in highly dense cities, but has no significant impact in cities with low to medium density and (ii) this air quality improvement is indeed mediated by the associated changes in commuting mode choices. Our findings provide strong empirical support for the hypothesis that the environmental impact of ride-hailing depends on its complementarity with public transit.

JEL Codes: L91, Q53, R4, R11

Key Words: Air pollution, commuting choice, satellite-based air quality data, staggered difference-in-differences, mediation analysis, ride-hailing, ride-sharing, transportation and environment

1. Introduction

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.*, 2016; Castillo, 2025) and from reducing information asymmetry (Liu *et al.*, 2018) and drivers' gain from Uber's compensation scheme (Angrist *et al.*, 2021) and from flexible workstyle (Chen *et al.*, 2019). Economists have also refuted some of the criticisms against these ride-hailing companies, and instead, find: Uber's entry increased public transit riderships (Hall *et al.*, 2018), reduced driving under the influence, fatal accidents, arrests for assault and disorderly conduct [Dills and Mulholland (2018), Anderson and Davis (2026)], and had no significant effects on taxi driver's labor supply although reducing their relative earnings by about 10 percent (Berger *et al.*, 2018). Against these economic benefits, however, a rising number of studies have also started to document the negative environmental impacts of ride-hailing services [Diao *et al.* (2021), Erhardt *et al.* (2019), Kong *et al.* (2020), Rayle *et al.* (2016), Tarduno (2021), Krishnamurthy and Ngo (2024)].

This manuscript seeks to delineate and empirically address an important yet unresolved question in the literature: *Is ride-hailing good or bad for the environment after all?* There is already abundant evidence showing that ride-hailing increases congestion on roads and highways, particularly in urban areas [Erhardt *et al.* (2019) and Tarduno (2021)]. There is also credible evidence establishing the causal relationship between this induced road congestion and an increase in ambient air pollution (Krishnamurthy and Ngo, 2024). This line of research is thought to confirm a behavioral pathway through which ride-hailing may increase road traffic and air pollution. This pathway is widely recognized in both the academic and non-academic literature. For example, an engineering study estimates that 'deadheading' between hired rides contributes 47% more emissions per trip while 'displacing' miles that might have been served by mass transit or other low-emission travel modes adds 69% more emissions per trip (Union of Concerned Scientists, 2020).

However, establishing the causal links along this pathway is hardly sufficient to answer the aforementioned question since, somewhat puzzlingly, it is also widely recognized that overall ambient air quality tends to improve, rather than deteriorate, after ride-hailing entry in many U.S. cities (Kim and Sarmiento, 2021; Konishi, 2024). This puzzle is, in fact, also documented by Krishnamurthy and Ngo (2024). They find that in California, weekday freeway traffic and $PM_{2.5}$ concentrations decrease in the "average county" upon Uber's entry, despite that freeway traffic and $PM_{2.5}$ concentrations increase during the evening rush hours in the populated counties. These findings, if true, suggest that there must be another counteracting pathway through which ride-hailing may decrease road traffic and air pollution. Hall *et al.* (2018) provide an important empirical finding that seems to shed light on such an alternative pathway: i.e., ride-hailing entry tends to increase the complementary use of public transit, particularly in large US cities. Indeed, it is straightforward to construct a numerical example of a monocentric city in which the combined use of ride-hailing may reduce *aggregate* road traffic and air pollution, while simultaneously increasing both in the city's urban core (see **Appendix B** for an illustrative example). The question is whether this latter pathway indeed exists empirically, and, if so, whether it is substantial enough to outweigh the former pathway in the average city entered.

Our goal in this paper is to investigate this question, empirically disentangling the chain of causal relationships along the latter pathway: i.e., between ride-hailing entry, commuting patterns, and overall air pollution. To do so, we proceed in two steps. Our first step is to credibly estimate the causal effect of ride-hailing on ambient air quality, exploiting the variation in ride-hailing entry over time and across cities in the staggered difference-in-differences (DD) design. Our basic empirical strategy is to estimate the Two-way Fixed Effect (TWFE) regression on monthly observations of ambient air pollution using a panel of 348 MSAs during the 9-year period, 2010-2018. The objective here is to replicate Kim and Sarmiento (2021) in a manner that is more credible and coherent to our goal. This step's empirical strategy is analogous to previous studies [Berger *et al.* (2018), Hall *et al.* (2018), Li *et al.* (2021), Ward *et al.* (2019), Kim and Sarmiento (2021)], but we take several new approaches to strengthen the identification of the estimand of interest.

First, we construct a new measure of ride-hailing entry into a MSA boundary, using Google's keyword search trends for both Uber and Lyft (hereafter, we call this "*de facto* entry measure" for brevity). This measure of entry has broader coverage than Uber's official launch data, is better aligned with the onset of locally salient and behaviorally relevant ride-hailing presence, 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 2 and 3**. Second, we use satellite-based nitrogen dioxides (NO_2) concentration data for improved identifiability of ride-hailing's impacts on *transport-related* air pollution. We specifically avoid using U.S. Environmental Protection Agency (EPA)'s monitoring-based data on NO_2 and other pollutants for various reasons. Monitoring sites are spatially unevenly located, are not necessarily located in high pollution areas (Fowlie *et al.*,2019), and change locations over time, making it difficult to compare data consistently over time. There may be 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). Using the satellite-based NO_2 data help us avoid these issues and track transport-related air pollution at the spatially disaggregate level consistently over time and across cities. Third, we use MSA-level population density as a sufficient statistic for unobservable trends in transport-related behavioral change, and estimate the heterogeneous dynamic treatment effects on each subsample of MSAs stratified by population density quartile, applying the de Chaisemartin-D'Haultfoeulle (dCDH, 2020; 2024) estimator.[1] This helps us avoid the contamination bias coming either from heterogeneous treatment effects or from the possibility that ride-hailing entry may occur in cities exactly when and where residents' transport-related behavior is expected to change.

Our second step is to credibly estimate the mediating effect of changes in modes of commuting to work on ambient air pollution, using household-level data from the American Community Survey (ACS) [this is also called as the "causal mediation effect" or the "indirect effect" in the mediation analysis literature (Imai, Keele, and Tingley, 2010; Imai, Keele, and Yamamoto, 2010; Huber, 2019)]. Because the survey data are available only yearly, we forgo sharp identification based on the timing of *de facto* entry. Instead, we exploit the within- and cross-city variation at the household and PUMA (public use microdata area) level, using observations in 2012 and 2016 only. These two years are well suited to serve as control and treatment periods because ride-hailing was not yet salient in 2012 in our sample, whereas by 2016 it had become behaviorally relevant in virtually all entered cities. Estimating the effect of ride-hailing entry on commuting modes is rather straightforward — we run the TWFE-DD regressions of yearly household-level commuting mode indicators on *de facto* entry, exploiting the MSA-level treatment-control structures. On the other hand, estimating the mediating effect of commuting modes on ambient air quality is quite challenging because we need exogenous sources of variation in commuting modes that are unrelated to ride-hailing entry and do not directly affect ambient air quality (Frölich and Huber, 2017; Huber, 2019). To construct such instruments, we combine Uber's official entry data with geography-based instruments that are widely used in the empirical economic geography literature [Baum-Snow (2007), Duranton and Turner (2011, 2012), Faber (2014), Redding and Turner (2015)] in the spirit of Bartik shift-share design. Specifically, we use Uber's official entry dates to construct a leave-one-out regional average exposure duration measure, which captures broader region-wide diffusion of ride-hailing while excluding the city's own contribution. We interact this timing component with the geography-based instruments, adjusting for the size of each PUMA. Intuitively, such instruments create counterfactual exposure duration at the

[1] Alternatively, we may use the Callaway-Sant'Anna (CS) estimator, which is theoretically identical with the dCDH estimator in the absence of covariates. As we explain in **Section 4**, we absorb the influence of all time-varying covariates using the imputation method (Caetano *et al.*, 2023). Hence, we should obtain the similar estimates, if not the same, using either estimator.

PUMA level to counterfactual transportation-network structure, which predict the timing and location of commuting mode changes while being less mechanically tied to ride-hailing's realized city-level penetration.

Using these empirical approaches, we obtain two sets of results corresponding to the steps explained above. First, we find robust evidence that ride-hailing entry tends to decrease ambient NO_2 concentrations (in terms of both monthly mean and maximum) in MSAs with high population density. The estimates for the highest density MSA quartile are always negative, have relatively small standard errors, and range from -0.032 log points to -0.044 log points using monthly means and -0.049 log points to -0.059 log points using monthly maximums. Assuming a linear relationship between satellite-based and EPA monitoring records, these impacts roughly translate into the reductions of monthly mean NO_2 concentrations by 0.50-0.69 ppb and monthly maximums by 4.63-5.57 ppb. The estimated impacts are slightly larger when using urban-area boundaries than suburban-area or non-urban-area boundaries, and also larger when using monthly maximums than using monthly means. This suggests that ride-hailing tends to reduce high-pollution incidence in large MSAs. We also find that ride-hailing entry has no significant effect on ambient NO_2 concentrations in low- and medium-density MSAs. We also confirm these findings in the event study of dynamic treatment effects using the dCDH estimator. Our event study suggests that ride-hailing tends to reduce ambient NO_2 concentrations only for the highest density MSA quartile and the magnitude of its impact gradually increases over time, although the estimates are somewhat imprecise due to the small sample size. For other density quartiles, we do not see any visible sign of ride-hailing's impact.

Second, we find evidence that the air-quality impact of ride-hailing indeed arises, at least in part, from its mediating effect through changes in commuting patterns. On average, ride-hailing entry is estimated to decrease the probability of private car commuting to work by 0.84 ppt and increase that of public transit commuting by 0.16 ppt. The impacts are largest in the highest density MSAs where ride-hailing entry is estimated to decrease NO_2 concentrations. More importantly, our IV estimates indicate that a 1-unit increase in the density-adjusted share of car commuters would increase NO_2 concentrations by 0.013 log points ($\approx$ 1.3%) whereas that of public transit commuters would decrease NO_2 concentrations by 0.007 log points ($\approx$ 0.7%). To place these numbers in economic scale, the IV estimates imply that if additional one percent of households were to switch their commuting modes from private cars to public transit in the average community, the monthly average NO_2 concentration would decline by approximately 0.30%. Furthermore, we combine these estimates to calculate the indirect estimates of ride-hailing's air quality impact implied by the causal mediation analysis. The results reveal that the magnitudes and directions of the estimates derived from the causal mediation analysis closely align with the direct estimates of the effect of ride-hailing entry on ambient air pollution, although the predicted impacts vary substantially with population density. Taken together, we conclude that the chain of causal relationships indeed exists between ride-hailing entry, commuting mode changes, and ambient air quality.

Our work complements several vibrant areas of research: (a) empirical studies that estimate the causal effects of ride hailing on various economic outcomes [Anderson and Davis (2026), Angrist *et al.* (2021), Berger *et al.* (2018), Chen *et al.* (2019), Cohen *et al.* (2016), Dills and Mulholland (2018), Hall *et al.* (2018), Liu *et al.* (2018), Tarduno (2021), Krishnamurthy and Ngo (2024)], (b) a large body of literature that examines pollution- or congestion-relief effect of public transportation infrastructures [Chen-Whalley (2012), Li *et al.* (2019), Gendron-Carrier *et al.* (2022), Gu *et al.* (2021), other papers cited in Anas and Lindsey (2011)], and (c) the economic studies that structurally investigate the general equilibrium impacts of ride-hailing into the taxi industry [Buchholz (2022), Fréchette *et al.* (2019), Hall *et al.* (2020), Rosaia (2025)]. Of these, our work is most closely related to Krishnamurthy and Ngo (2024), Kim and Sarmiento (2021), and Tarduno (2021). Krishnamurthy and Ngo (2024) use hourly freeway traffic and daily air pollution data from California and apply a DD design exploiting the staggered rollouts of ride-hailing entry at the county-level. Kim and Sarmiento (2021) use the Callaway-Sant'Anna estimator in an empirical design similar to ours and find that Uber's entry is estimated to improve air quality (as measured in the Air Quality Index and ground-level ozone), particularly during the summer when bad air quality episodes are expected. We emphasize here that, while our findings are consistent with theirs, our work is fundamentally different, not only in the study's scope and objectives, but also in design and implementation of empirical strategies. In particular, we explore the alternative behavioral pathway through which ride-hailing may decrease road traffic and air pollution, and find credible evidence in support for the causal links between ride-hailing entry, commuting pattern changes, and a reduction in overall air pollution. Some of these studies have also explored this pathway, but did not formally conduct the causal mediation analysis like ours (see also **Appendix A** for a more complete review of the literature on the environmental concerns of ride-hailing services). Lastly, Tarduno (2021) estimates the impact of the exit of ride-hailing services on traffic by exploiting the abrupt departure of Uber and Lyft in 2016, prompted by the enforcement of a fingerprinting requirement in Austin, Texas. While he finds that the *exit* of ride-hailing services reduced congestion — as measured by traffic speeds — by approximately 2.3% on average, the effects were spatially heterogeneous: congestion increased mostly in outer-city segments but decreased in inner-city areas. The implications of these heterogeneous congestion effects for overall urban air pollution remain ambiguous, as do the underlying causes of such spatial variation. Our paper provides a conceptual framework to explain the mechanisms driving these heterogeneous responses, employs both descriptive and causal mediation analyses, and presents evidence of a pollution-reducing effect linked to this mechanism.

2. Background and Motivating Facts

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, bringing "transportation innovations" there.

The ultimate goal of our manuscript is to address an important, unsettled controversy: *Q1. Do ride-hailing services decrease or increase transport-related pollution?* The motivation for our work is best illustrated by **Figure 1**, which displays the time series of three datasets used in this study. How we construct the data is explained in detail in **Section 3**. For an illustrative purpose, we plot the data only for the city of Chicago, IL.

Panel A plots the time trends of monthly taxi trip miles (blue) and ride-hailing (taxi + rideshare (TNC)) trip miles (red). The rideshare trip data are available only after November 2018, and hence, between 2013 and 2018, we plot the predictions obtained from a third-degree polynomial regression of rideshare trip miles on Google Trends Index (the details of the index are in **Section 3**). The figure shows that ride-hailing trip miles in 2016 and later are an order of magnitude larger than taxi trip miles in 2013, a period prior to ride-hailing substantial market penetration. On the other hand, the figure also indicates the official entry dates of Uber, Lyft, and UberX into Chicago. Those are the dates used in previous empirical studies for causal inference. It is clear that use of these entry dates is unlikely to achieve the kind of counterfactual comparison we wish to make: The air quality level in the state in which ride-hailing services are behaviorally relevant versus the state in which they are not.

We construct an alternative Google Trends-based measure of entry and use it in combination with other empirical strategies to alleviate this problem.

Panel B shows the time series of monthly MSA-level averages of ambient NO_2 concentrations in Chicago using the satellite-based data (light green area) and EPA's monitoring data (red circles). Satellite-based data allow us to calculate the area-wide averages for non-urban, suburban, and urban areas. The dashed line indicates the suburban NO_2 concentration levels, with the light green band bracketed by the urban and the non-urban NO_2 trends. For this figure only, the satellite-based data are converted into comparable units in part per billion (ppb) by fitting a linear regression of satellite-based data on EPA monitoring data (see **Appendix D** for more details). The figure demonstrates that satellite-based data and monitoring data show similarly cyclical NO_2 trends, yet the former shows far more variability with higher and lower peaks. This is because the monitoring data are available only for a limited number of monitoring stations, most of which are located near the city center or the major highways. Another important take-away is that there seems no visible sign of change in NO_2 concentrations that would correspond to any of the official entry dates. Furthermore, the non-seasonal cyclicality of ambient air pollution tends to obscure the potential impact of ride-hailing activity, and hence, makes it difficult to rely on sharp identification strategies such as regression discontinuity in time or space.

Panel C plots the time trends in the MSA-level commuter shares by mode of transportation to work, using the household-level data from the ACS. We see an increasing trend in the share of commuters using public transit and other modes, alongside a declining share of those commuting by private car. Two important questions arise immediately from the figure: *Q2. Are these changes in commuting mode shares causally related to the use of ride-hailing? Q3. If yes, are the changes in ambient* NO_2 *concentrations also causally related to these changes in commuting modes?*

The latter question is quite important in light of the prior literature. Kim and Sarmiento (2021) find that use of ride-hailing may causally reduce ambient air pollution, but do not fully establish the underlying mechanism. Hall *et al.* (2018) and Krishnamurthy and Ngo (2024) show that public transit riderships increase upon Uber's entry in large cities, yet *do not* show that this ridership change occurs as the commuting mode changes among commuters, nor that air pollution changes as a result of the ridership change. Krishnamurthy and Ngo (2024) also explore several other pathways, and suggest that cars used for ride-hailing services may be cleaner, potentially displacing dirtier cars. Tarduno (2021) finds that the *exit* of ride-hailing services decreased congestion primarily in inner-city road segments but increased it in outer-city segments; however, he does not explore the mechanisms underlying these heterogeneous responses or their implications for air pollution. The figure also illustrates the empirical challenges we face in addressing these questions. Commuting mode data are available only on a yearly basis, and their yearly trends are most likely influenced by time-varying confounds. To alleviate the problem, we use within- and cross-city geographic variation in both satellite-based NO_2 concentrations and commuter shares.

Figure 2 helps us further elaborate our empirical context and explain why we have a reason to suspect ride-hailing might have induced a substantive change in commuting patterns. **Panel A** plots the average distance per trip of all taxi rides during the morning and evening rush hours against their pickup location bins measured in distance to the central business district (CBD) and the distribution of taxi trip volumes per month over the pickup locations, in 2014 (before ride-hailing's full market penetration) within the community area boundary of Chicago. The whiskers represent the 25th and the 75 percentiles of trip distance and the 45-degree line is drawn for ease of interpretation. **Panel B** plots the same using all rideshare (TNC) trips during the morning and evening rush hours in 2019 (after ride-hailing's full market penetration). The idea here is that if the hired ride is used to commute all the way to CBD from the pickup location, then the trip distance (shown in the vertical axis) should be greater than the distance to CBD (shown in the horizontal axis). For instance, roughly 75% of ride-hailing trips originating from O'Hare Airport lie above the 45-degree line, implying that customers on these trips are likely to be traveling directly to CBD. The figure demonstrates that in 2014, only commuters near CBD use ride-hailing (taxi) during the rush hours, and when they do, a majority of them ride for very short distance. In 2019, in contrast, commuters outside CBD use ride-hailing (TNC) during the rush hours, and they ride for much longer distance. The average trip distance gets smaller than the distance to CBD, however. Roughly 75% of weekday rush-hour rides that originate from 5 miles away or further from CBD do not directly travel to CBD. This pattern is consistent with a complementary relationship between ride-hailing and public transit services, as suggested by Hall *et al.*, although it may also reflect increased intra-suburban travel facilitated by ride-hailing rather than complementarity with transit.

The figure also explains why ride-hailing may decrease or increase transport-related air pollution, signifying the importance of our empirical questions. On one hand, the increase in ride-hailing trip volumes should increase air pollution, holding all else constant. This is line with the critics' argument that ride-hailing can add road traffic either because of 'dead-heading' (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. On the other hand, if a majority of commuters use ride-hailing in combination with other public transit, then ride-hailing can potentially reduce private driving *more than* simply replaces it. This argument aligns with the proponents of Uber and Lyft claiming 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. If this were the case, air pollution may decrease at least during the rush hours.[2] In **Appendix B**, we use a simple example of a monocentric city to clarify (i) the condition under which ride-hailing can reduce private driving more than simply replace it while increasing congestion in city centers, (ii) both demanders and suppliers of ride-hailing are important in understanding its air quality effects, and (iii) the economic logic underlying the interpretation of the estimates from our mediation analysis.

3. Data

To achieve our goal, we run three types of regressions: (1) air quality on ride-hailing, (2) commuting modes on ride-hailing, and (3) air quality on commuting modes. Our study sample consists of 348 U.S. metropolitan statistical areas (MSAs) using the 2010 Census boundaries, but the level of aggregation differs across regressions to make most of all available variations in data. We use monthly panel data at the MSA level for (1), yearly panel data at the household level for (2), and yearly panel data at the PUMA level for (3). The following sets of data and variables are used for these regressions.

A. Ride-hailing Entry: The chronology of the ride-hailing industry makes it difficult to define "entry" of ride-hailing service in a city. First, earlier studies often rely on *reported* Uber/Lyft entry dates compiled from newspapers, company blogs, and other public sources, which may be incomplete or inaccurate. Second, Uber and Lyft entered different cities in different timings with varying degrees of market presence, while the presence of either platform may be sufficient to affect commuting patterns. Third, official launch dates need not coincide with the time at which ride-hailing becomes locally salient and behaviorally relevant. Earlier studies (Hall *et al.*, 2018; Krishnamurthy and Ngo, 2024) do acknowledge these concerns, and address them by including sufficient post-entry data to ensure that full penetration is reached in all markets.

Our approach to overcoming these challenges is, instead, to rely on Google Trends data and construct a measure of the onset of locally salient and behaviorally relevant ride-hailing presence. The measure is intended to capture when ride-hailing begins to matter for behavioral change, not the exact timing of formal launch or widespread market penetration. For

[2] Air pollution data are often recorded as daily averages or daily maximums, and daily air pollution peaks occur during morning and evening rush hours when vehicle transportation volumes are the highest. Hence, if we were to find the effect of ride-hailing on transport-induced air pollution, it is most likely to come from changes in commuting patterns during these rush hours.

brevity, we refer to this measure as *de facto* entry, although it should not be interpreted as a precise measure of entry or market penetration. Specifically, we construct this measure based on two normalized Google Trends indices. For each MSA in our sample, we obtain search trend indices from January 2010 to December 2018, using "Uber" and "Lyft" as keyword entries.[3] We then define entry if the maximum of the two indices exceeds a certain threshold. That is, for each MSA i and for each month t,

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

For all subsequent analyses, we use the cutoff value of $c = 0.2$. We tested several values for the cutoff, but eventually chose the cutoff value of 0.2 to match the ideal of our quasi-experimental design. That is, the cutoff value must be such that below the cutoff, there is virtually no Uber/Lyft search 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. **Appendix C** provides a more detailed explanation of the steps taken to construct this measure.

Panel (a) of Figure 3 compares the distribution of our *de facto* entry dates constructed from this measure against that of (i) *reported* entry dates of UberX, taken from Hall *et al.* (2018) and (ii) *official* Uber launch dates obtained from Uber Inc.[4] Roughly, Uber's official launch date for each MSA is defined as the month after which 25 or more "active" drivers, who complete at least 4 trips per month, continue to exist in that MSA. The figure shows that our Google Trends-based measure differs systematically from both reported and official entry dates, consistent with our *a priori* expectations. **Section 4.D** and **Appendix C** discuss the interpretation of these discrepancies, carefully delineate the settings in which our measure appears strongest and those in which alternative interpretations remain more plausible, and report the robustness of the main results to a more conservative entry measure.

Panel (b) of Figure 3 compares the event-study estimates of the effect of entry on the normalized Google Trends index using the three entry measures.[5] Four points are worth noting. First, the estimated effects of reported and official entry remain quite small over the two-year window after entry and are even smaller than the cutoff value ($c = 0.2$) used to define *de facto* entry. Thus, if reported entry were used as the treatment variable, the estimated treatment effect would reflect only a low level of Uber/Lyft activity. Second, there

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

[4]Our special thanks go to Michiko Namazu, a data scientist at Uber Inc., who kindly shared the official launch data with us. We also thank our referees and the editor for encouraging a fuller discussion of **Figure 3** in **Section 4.D** and **Appendix C**.

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

is a discontinuous jump at the *de facto* entry timing, suggesting that the Google Trends index rises sharply around the cutoff. We view this as evidence that the cutoff is reasonably calibrated to mimic the quasi-experimental design. Third, the estimated effects of *de facto* entry continue to increase over the two-year window, implying that the treatment effect on air pollution is likely to be dynamic as well. Fourth, there is no visible sign of a violation of the no-anticipation assumption: the pre-entry estimates are close to zero and relatively precise.

B. Ambient Air Quality: Our second data source is the satellite-based NO_2 data from the National Aeronautics and Space Administration (NASA)'s Goddard Earth Sciences (GES) Data and Information Services Center (DISC). GES-DISC provides level-3 daily total column NO_2 in molecules per cm^2 on the 0.25×0.25 degree global grids. We focus on ambient NO_2 pollution because it is a key transport-related pollutant and more closely tracks emissions from on-road vehicles than broader pollutants such as $PM_{2.5}$.[6] We use the satellite-based data because they provide broad spatial coverage, within-city granularity, and intertemporal consistency, whereas EPA monitoring data are available only for a limited subset of MSAs, are unevenly located within cities, exhibit substantial attrition over time, and may also reflect non-random measurement due to strategic compliance behavior (Grainger and Schreiber, 2019; Grainger et al., 2019; Zou, 2021). For illustration, **Figure 4** plots the satellite-based NO_2 grid data together with EPA monitoring-site locations for two example cities. The figure shows that monitoring sites are sparse, concentrated near city centers, and do not necessarily coincide with high-pollution areas, while elevated NO_2 incidence can extend beyond urban cores into suburban and non-urban areas. This is important in our setting because ride-hailing may have heterogeneous air-quality effects across within-city subareas, and the satellite-based data allow us to construct area-weighted monthly means and maxima separately for urban, suburban, and non-urban areas within each MSA. Consistent with earlier studies, the satellite-based measures are strongly correlated with EPA monitoring data where overlap exists. **Appendix D** provides a fuller comparison and discusses the extent of bias in the EPA data.[7]

[6] NO_2 is a leading cause of respiratory diseases such as asthma, a known precursor to ozone, and is commonly used as an indicator for the broader group of nitrogen oxides (NO_x). Road transportation accounts for roughly 36% of total NO_2 emissions (larger shares in urban areas), but accounts for only 2-4% of $PM_{2.5}$, PM_{10}, and VOC emissions in the U.S, according to EPA's National Emissions Inventory. We, therefore, expect the air-quality effect of ride-hailing to appear most clearly in NO_2. We do not use carbon monoxide (CO) because we only have access to the satellite-based CO data on much less granular level.

[7] **Appendix D** shows that the EPA data suffer from substantial sampling problems even in large MSAs. The number of monitoring sites differs markedly across cities and often change within cities over time, which can mechanically shift average pollution levels. The extent of this problem is highly uneven across cities: in some places monitoring locations change frequently, while they remain relatively stable in others such as some California cities, as noted by Krishnamurthy and Ngo (2024).

C. Commuting Mode: We use yearly household-level commuting-mode data from the American Community Survey (ACS) to study the behavioral pathway between ride-hailing entry and NO_2 concentration. In the ACS, respondents report the main mode they used to commute to work during the week before the survey; if more than one mode is used, they report the one used for most of the distance. We use these responses to construct an indicator $I_{it}(\kappa)$ for household i's commuting mode κ to work in year t, which serves as our mediator. For ease of interpretation, we group the 12 categories of transportation modes into three broader modes: private car ($I(c)$), public transit ($I(pt)$), and other modes ($I(o)$). The first two account for 91% of commuting choices in our sample (85.6% and 4.8%, respectively). The residual category includes walking, bicycling, working at home, and taxicab or motorcycle use. Because the ACS does not separately identify ride-hailing use, we do not attempt to infer ride-hailing as a primary commuting mode. The ACS microdata are repeated cross-sections rather than a household panel, so our panel structure is at the PUMA level. We, therefore, supplement the analysis with household-level demographic controls such as education, income, and race.

D. Instruments: For the mediation analysis, we use an IV approach in the spirit of Bartik shift-share design, which combines official Uber rollout profile with three geography-based instruments that are widely used in the empirical economic geography literature. The former component is constructed from Uber's official entry dates obtained from Uber Inc., while the latter component uses the planned-route, historical-route, and inconsequential-unit instruments of Baum-Snow (2007), Duranton and Turner (2011, 2012), and Faber (2014). These geography-based variables are defined at the MSA level. **Section 5** explains how we combine these two components to construct our final IVs and discusses the identifying assumptions. **Appendix G** provides the details and maps of the geography-based instruments, and **Appendix I** reports first-stage diagnostics and robustness to alternative instruments.

E. Control Variables: We use a few other sources to construct time-varying controls. These include monthly temperature and wind data from EPA AirData, regional gasoline prices from the North American Electric Reliability Corporation (NERC), population density from the U.S. Census Bureau, and county-level nonattainment status for EPA criteria pollutants from the Greenbook database. We include the nonattainment indicators because local compliance efforts under the national air quality standards may affect the trends in NO_2 through regulatory responses to related pollutants such as $PM_{2.5}$ and O_3.[8]

[8] There is a potential "bad controls" concern with these nonattainment indicators. On one hand, failure to control for them could bias the estimated effect of ride-hailing entry upward if regulatory responses coincide with entry timing. On the other hand, if ride-hailing itself affects ambient NO_2, and thereby indirectly affects

F. Chicago Taxi and Rideshare Trip Data: For illustrative purposes, we use administrative data on taxi and transportation network company (TNC) trips, as well as public transit ridership, obtained from the City of Chicago's Data Portal (data.cityofchicago.org). The data include complete records of taxi trips from January 2013 onward and TNC trip records from November 2018, covering all trips taken within Chicago's community area boundaries. For each trip, the dataset reports fare amounts, trip distance, and pickup and drop-off times and locations. We use these data to construct monthly time series of taxi and ride-hailing trip miles (**Figure 1-A**), the geographic distribution of taxi and ride-hailing trips during peak travel hours (**Figure 2**), and the hourly distribution of taxi and ride-hailing trips (**Figures A2** and **A3** in **Appendix B**). These data are used solely for illustrative purposes and are not used in the formal empirical analysis.

Table 1 provides a statistical overview of the main outcome and pre-treatment variables by treatment status. The earliest *de facto* entry occurs in 2013, while 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, satellite-based NO_2 data suggest that this may be due to systematic bias in the monitoring data. EPA monitoring data are available for about 22–27% of observations for cities where Uber/Lyft enter before 2015, but for only 5–8% of observations for cities where Uber/Lyft entry never occurs or occurs after 2015. Using the satellite-based data, differences in NO_2 concentrations across cities by entry status become less obvious.

4. Air Quality Impact of Ride-hailing

4.A. Estimation and Identification Strategy

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

$PM_{2.5}$ or O_3 attainment status, then including these controls could attenuate the estimated entry effect. We view the latter concern as limited, since transport-related emissions account for only a small share of PM and VOC emissions, so any resulting bias is likely to be small even if it exists.

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

As a basis of our analysis, we start with the standard two-way fixed effect (TWFE) specification of the staggered DD regression, following Hall *et al.* (2018) and Berger *et al.* (2018). For an outcome variable $y^k_{s,ctm}$ in year t and month m of city (MSA) c in s-th quartile of MSA-level population density (using geographic boundary k), we specify our TWFE regression as:

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

where D_{ctm} is our treatment variable and equals 1 if ride-hailing service enters/exists in MSA c and 0 otherwise in period tm, δ_s is the heterogenous treatment effect parameter corresponding to each $s \in \mathcal{S} \equiv \{1, 2, 3, 4\}$, X_{ctm} is a vector of (exogenous) time-varying covariates, α_c and λ_{tm} are MSA and year/month fixed effects, and $\theta_c(t)$ is a MSA-specific linear time trend. Our main outcome variables of interest are ambient levels of nitrogen oxides (NO_2) concentrations in logged terms. 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.[9] 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 quartile 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 reporting

[9] We use temperature and wind data from the monitoring records available at the EPA's AirData. As we discuss in **Section 3**, 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.

errors. More importantly, in some cities, reported entry occurs well before the ride-hailing service becomes sufficiently salient. 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 Google Trends-based measure of entry (see our discussion in S**ection 3**).

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 NASA's satellite-based NO_2 concentration data, which are available at relatively high resolutions ($0.25° \times 0.25°$ girded cells) consistently throughout our study period (see our discussions in S**ection 3** and in **Appendix D**).

Third, although NO_2 has a known distance-decay relationship (Cape *et al.*, 2004; EPA, 2008), it may still travel far distance and affect ambient concentration levels several miles away from its emissions sources (Su *et al.*, 2009).[10] This implies that the violation of STUVA (or no-spillover) condition may occur in several ways, particularly between neighboring MSAs. To avoid such spillover effects, we delineate urban, suburban, and non-urban boundaries within each MSA, calculate the monthly NO_2 concentration statistics for each of these subarea boundaries $k \in \{\text{urban, suburban, non-urban}\}$, and run a separate regression for each k in eq. (1). This allows us to compare, for example, urban-area NO_2 concentration of entry city against urban-area NO_2 of no-entry city. Because urban areas are quite far apart from each other, this minimizes the risk of STUVA violation. Furthermore, our discussion in **Section 2** suggests that ride-hailing entry may have different air-pollution impacts on different subareas of cities. Since we expect some spatial spillovers across subareas of cities, how the treatment effects differ over subareas of cities is also an important empirical question.

Besides these empirical challenges, we have an important identification challenge. The major threat to identification in eq. (1) comes from the violation of parallel trends in unob-

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

servables. As discussed in Hall *et al.* (2018), ride-hailing services mostly enter cities in the rank order of population size (or population density). We also confirm this point more formally in **Appendix F**. 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). To overcome this challenge, we strengthen our identification in the following manner.

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

Assumption 1. *Conditional on exogenous (time-varying) covariates $\Omega_{c,s}$, parallel trends in the absence of treatment hold for all MSAs in the same population density cohort $s \in \mathcal{S}$: i.e.,*

$$E[y_{c,\tau}(0) - y_{c,\tau-1}(0)|D_{c\tau} = 1, \Omega_{c,s}] = E[y_{c,\tau}(0) - y_{c,\tau-1}(0)|D_{c\tau} = 0, \Omega_{c,s}],$$

for cities within each population density quartile s for all periods τ.

Note that in the staggered DD setup, parallel trends need to hold for each timing group: i.e., outcomes in the absence of treatment 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 fully interact the population density quartile dummies with the treatment variable. This is the approach we take with our TWFE specification in eq. (1). The second is to estimate the treatment effect parameter separately on each density-quartile subsample. This approach is taken in our event-study estimation below. The first approach compares the outcome trends of the treated cities against those of the not-yet-treated within the same density cohort s, controlling for the set of included covariates in eq. (1). The second approach compares the outcome trends between the treated cities versus the not-yet-treated cities within the same density cohort s, not only controlling for the set of included covariates but also purging out the influence of all cohort-level unobservables $\omega_{s,\tau}$ that may vary over time systematically differently across density cohorts. Therefore, the second approach is more robust to misspecification 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 F**). 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 quartile 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 (2025)]. Furthermore, our specification also helps us explicitly account for heterogeneous treatment effects of ride-hailing entry to vary by (pre-treatment) population density. This is important because we expect the complementarity between ride-hailing service and public transit to be greater in cities with density (both public transit density as well as proximity to each other are important). Słoczyński (2022) 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.

4.B. Event Study

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

Gardner (2022) and Caetano *et al.* (2023): we can use pre-treatment observations to purge out the effect of (exogenous) time-varying confounders. Specifically, we take the following steps:

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

$$y_{ctm}^k = \alpha_c + \lambda_{tm} + X_{ctm}'\gamma + \theta_c(t) + \epsilon_{ctm} \tag{2}$$

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

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

3. Apply the following dCDH/CS estimator to the estimated residuals, separately for each subarea k and each density-cohort subsample s (not pooling all MSAs): For MSAs that receive treatment in g-th period, estimate the average treatment effect on the treated (ATT) for τ-th period as[11]

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

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

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

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

they are correlated with treatment assignment. We avoid the bias by estimating the parameters using pre-treatment observations. Third, this approach explicitly makes use of the **assumption 1,** and estimate ATTs using only observations in the same density cohort. Hence, our approach compares only 'similar units' in terms of the pre-treatment population density.[12]

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.C. Estimation Results

We start by presenting the estimation results from the TWFE regression of eq. (1). 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 O_3 and $PM_{2.5}$. Inclusion of endogenous time-varying controls can potentially bias our estimates since they can be correlated with unobservables in the estimating equation. Hence, we avoid use of such controls. Standard errors are clustered at the MSA level.

In both tables, there is a tendency that the estimated impacts tend towards positive values with more controls, suggesting that our estimates may be biased towards negative values. Hence, we focus on the results with full controls. In **Table 2-A**, we see that *de facto* ride-hailing entry is estimated to decrease ambient concentration levels of NO_2 for cities in the highest population density quartile. The estimates are statistically significant, and range from -0.032 log points (in non-urban areas) to -0.034 log points (in urban areas). If we use a linear transformation evaluated at the mean concentration levels, these estimates imply that *de facto* ride-hailing entry reduces NO_2 concentrations by 0.50-0.53 ppb for these

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

cities (roughly 0.06 standard deviation decrease or 5.3-5.8 percent decrease relative to the EPA sample mean). Interestingly, the estimates get smaller in magnitude for less densely populated cities, and eventually turn positive for cities in the lowest density quartile. The results seem consistent with our story — In cities with sufficient public transit networks, a combined use of ride-hailing with public transit can displace use of private cars, reducing overall vehicle emissions, but in cities without such public transit networks, ride-hailing may simply induce more driving, increasing overall vehicle emissions. However, whether such a causal pathway indeed exists (and how large the magnitude is if it does) is an empirical question, which we explore more fully in the next section.

We turn to the results on monthly maximums in **Table 2-B**. The results are generally consistent with those of **Table 2-A.** That is, the estimates are negative and statistically significant for cities in the highest density quartiles; the estimates get smaller for cities in the lower density cohorts, and turn positive for non-urban areas in the lowest density cities. There is one important difference, however. The estimated impacts are larger, more statistically significant, and range from -0.049 log points (in suburban areas) to -0.050 log points (in urban areas) (roughly 0.09 standard deviation decrease or 10.6-10.8 percent decrease relative to the EPA sample mean). 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, hours, or areas where pollution levels are high (i.e., heavy vehicle traffic are expected). This also explains the relatively large estimates on the mean concentration levels because the monthly means are highly sensitive to the maximum concentration levels. These results are also consistent with Kim and Sarmiento (2021), who finds that much of the air quality improvement comes from a decline in the number of bad air quality days in summer.

Next, we present the event study estimates from the dCDH estimator.[13] As explained in **Section 4.B**, we implement the dCDH estimator separately for each density quartile (rather than pooling all observations) after purging out the influence of exogenous time-varying covariates. Hence, the identification is much clearer here. By construction, this estimator uses only 'not-yet-treated' (or 'never treated') MSAs for each treatment timing within the same population density cohort as comparison units. As a result, the estimates are not contaminated from use of either already treated MSAs or other non-similar MSAs in different density cohorts.

Figure 5 plots the estimates of the ATTs in relative time to entry ($\tau = 0$). The first two graphs on the top panel plot the estimates pooling all MSAs. The other graphs plot the

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

estimates using a subsample of MSAs on each population density quartile. To avoid busy graphs, we only present the results using monthly maximum NO_2 except for **Panel A** on the pooled sample. For the highest density cohort, eventually all units get treated, and hence, we are able to estimate the ATTs only within the two-year window. Thus, for consistency, we use only not-yet-treated units as control units and estimate the ATTs over the two-year window for all density cohorts.

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

4.D. Discussion

Before turning to our mediation analysis, we discuss three issues that are important for interpreting the main results. First, we rely on the Google Trends indices to infer when ride-hailing becomes sufficiently salient and behaviorally relevant in the local transportation market. While we are confident about the care we have taken in constructing this measure, a genuine concern arises when comparing the distribution of *de facto* entry to the reported and official entry dates in **Figure 3-(a)**. Our *a priori* expectations are that (a) *de facto* entry lags reported and official entry in large cities where Uber engaged in early, limited rollouts; (b) *de facto* entry precedes reported and official entry in a subset of cities where Lyft penetrated prior to Uber; and (c) *de facto* entry captures broader coverage than reported or official entry in small and medium-sized cities. Case (c) can arise because Uber's official launch date is defined conservatively and may therefore understate the locally relevant availability of ride-hailing services, particularly in smaller cities. We also verified that all cities classified as having *de facto* entry had ride-hailing services available as of 2020, even though many are recorded as having no official Uber entry. **Figure 3-(a)** is broadly consistent with these expectations: *de facto* entry tends to lag administrative entry measures in early-entry cities, while it precedes them in later-entry or administratively unrecorded cities. At the same time, the figure shows that *de facto* entry precedes reported or official entry for a nontrivial share of cities, suggesting that our Google Trends-based measure may capture anticipation or salience in the absence of behaviorally relevant ride-hailing presence.

We address this concern in three ways. First, we construct an alternative *de facto* entry measure that excludes Lyft-related keyword searches and report the resulting distribution of entry timings by density quartile in **Figure A5** of **Appendix C**. The results show that the discrepancy between *de facto* and official entry timings is concentrated in smaller, lower-density cities, and that removing Lyft-related searches reduces — but does not eliminate — the gap between the two distributions, particularly in the highest density quartile. Second, we construct the same Uber-only *de facto* entry measure using only the largest 150 MSAs, ranked by 2010 population. For these cities, the cumulative distribution of *de facto* entry indeed lies below that of official Uber entry, consistent with the interpretation that the Google Trends-based measure more closely tracks behaviorally relevant ride-hailing presence in larger markets. Third, we refine our *de facto* entry measure by constraining it not to precede Uber's official entry date, and re-estimate the TWFE and event-study regressions using this refined measure. The results are largely robust: the estimates remain negative and statistically significant in the highest density quartile and mostly insignificant in the remaining quartiles, for both monthly mean and monthly maximum NO_2. The magnitudes are also similar for monthly means, though somewhat smaller for monthly maximum. Taken together, these results suggest that our Google Trends-based measure is a reasonable proxy for the onset of behaviorally relevant ride-hailing presence in large, early-entry cities. In smaller or later-entry cities, however, it may also reflect anticipation or salience, so the treatment should not be interpreted as market penetration there.

Second, the credibility of our results also depends on the quality of the satellite-based NO_2 data. We take a number of steps to validate these measures. **Appendix D** presents descriptive analyses demonstrating: (i) the satellite-based NO_2 data provide a valid substitute when and where EPA monitoring data are unavailable; (ii) EPA data on NO_2 and $PM_{2.5}$ may be missing non-randomly, even for large MSAs; and (iii) the satellite-based NO_2 data exhibit sufficient variation — including during pollution peaks in smaller cities — and closely track the cyclical patterns, levels, and variability of both monthly means and maximums observed in the EPA data. We also estimate TWFE regressions of ride-hailing entry on air quality using EPA monitoring data on NO_2 and $PM_{2.5}$ for the subset of cities where such data are available. As reported in **Appendix D**, the signs and magnitudes of these estimates are similar to our main results.

Third, we classify MSAs into quartiles based on population density and estimate the TWFE and event-study parameters separately by quartile, under the assumption that parallel trends in outcomes hold within each density cohort. Given our results mostly come from the highest density quartile, two important questions arise: (i) Which cities fall into this quartile, and are they transit-friendly? and (ii) Are our findings robust to alternative cohort definitions more directly linked to public transit availability/usage? **Appendix E** lists the 40 most densely populated MSAs. The list includes well-known transit-friendly cities — such as New York, San Francisco, Boston, Washington, D.C., Philadelphia, Chicago, Seattle, Newark, Portland, Baltimore, and Minneapolis-St. Paul — where rail and subway systems are in operation. However, the list also contains cities that are not traditionally transit-friendly. Thus, it is plausible that our results are primarily driven by these transit-friendly MSAs. To explore this point, we estimate three sets of TWFE regressions using alternative cohort definitions based on public transit characteristics: (i) MSA-level public transit network density (using data on bus networks as of 2021 and rail networks as of 2012), (ii) the average share of commuters using rail transit at the MSA level (as of 2010), and (iii) the average share of commuters using bus transit (also as of 2010). As shown in **Appendix E**, the main findings are robust to these alternative cohort definitions and further suggest that the pollution-reducing effect of ride-hailing is strongest in cities with higher pre-treatment rail transit use.

5. Mediation Analysis

5.A. Overview

So far, our results are consistent with the behavioral pathway outlined in **Section 2** (and **Appendix B**). We have yet to establish that the chain of causal relationships indeed exists along such a pathway. We first clarify our empirical object of interest, following the potential outcome framework for causal mediation analysis outlined in Imai *et al.* (2010), Imai *et al.* (2010), Huber (2019), and VanderWeele (2016). Let Y denote the city-wide air quality, D the ride-hailing's market penetration status, and M the mediator of interest, in our case, some measure of the city-wide commuting pattern. M_1 and M_0 denote the potential commuting pattern with and without the ride-hailing's market penetration, respectively. Given this framework, the mediation analysis decomposes the average treatment effect (ATE) into the average direct effect (ADE) (evaluated at $M = M_1$) defined as:

$$ADE \equiv E[Y|M_1, D = 1] - E[Y|M_1, D = 0], \tag{4}$$

and the average indirect effect (AIE) (evaluated at $D = 0$) defined as:

$$AIE \equiv E[Y|M_1, D = 0] - E[Y|M_0, D = 0]. \tag{5}$$

ADE is the change in city-wide air quality directly caused by ride-hailing services, holding the city-wide commuting pattern. AIE is the change in city-wide air quality caused by the change in city-wide commuting pattern, controlling for the direct effect of ride-hailing on air quality. By definition, $ATE = ADE + AIE$. We can also analogously define ADE holding M at $M = M_0$ in eq. (4) and AIE holding D at $D = 1$ in eq. (5). But we use the definitions above to keep our discussion and notation simple. We can empirically confirm the causal mediating mechanism if we can empirically reject the null of no mediation effect ($AIE = 0$) and the null of no treatment effect on the mediator variable ($E[M_1 - M_0] = 0$).

There is a rich literature discussing a variety of identification strategies for the causal mediation analysis. One approach is to follow Imai *et al.* (2010) and achieve identification via the sequential ignorability assumption. Another approach is to apply an instrumental variable (IV) method in the system of equations (see Frölich and Huber (2017) and Huber (2019) for an excellent review on this approach). We use a simple, yet empirically robust version of the latter approach. Our identification strategy is similar to those used in previous empirical studies such as Chen *et al.* (2018), Chen *et al.* (2019), and Powdthavee *et al.* (2015). Ours, however, differs from theirs in that we employ an identifying assumption just enough to credibly identify and estimate the empirical parameter of interest, not all parameters of the system of equations.

To see this, consider the following canonical linear equation system:

$$\begin{aligned} Y &= \alpha_1 + \beta_1 M + \delta_1 D + \lambda_1 (D \times M) + X'\gamma_1 + \epsilon_1, \\ M &= \alpha_2 + \delta_2 D + X'\gamma_2 + \epsilon_2. \end{aligned} \tag{6}$$

All structural parameters of this system is fully identified under the sequential ignorability assumption (Imai *et al.*, 2010) or partially under the exclusion restriction ($\delta_1 = 0$ and $\lambda_1 = 0$) given the ignorability of treatment D (Angrist *et al.*, 1996). Either way, the identification relies on very restrictive assumptions that are unlikely to hold in our empirical context. Note, however, that given this system of equations, AIE can be written as

$$AIE = \beta_1 \left\{ E[M|D = 1] - E[M|D = 0] \right\} = \beta_1 \delta_2. \tag{7}$$

Therefore, we do not need to identify all parameters of the system to test the causal mediation mechanism. All we need to do is to credibly identify and estimate β_1 and δ_2. Our identification argument for β_1 is that there exists an instrument Z such that Z is correlated with M, but is uncorrelated with D, and does not directly affect Y. Once we have such Z, we estimate eq. (6) using the two-stage least squares (2SLS) estimator, directly controlling for other terms using fixed effects and covariates. For δ_2, we essentially make use of the same identifying assumption as in **Section 4** — parallel trends in M for each density cohort. Lastly, we emphasize here that our estimand of interest, AIE, is not meant to capture the direct effect δ_1 nor the interaction effect λ_1. Hence, we expect the estimate $\hat{\beta}_1\hat{\delta}_2$ to differ from the overall air quality impact of ride-hailing estimated in **Section 4**. We discuss this point more fully when discussing our results below.

5.B. Estimation and Identification Strategy

To operationalize the mediation analysis, we start by condensing our staggered DD design into a simpler 2-by-2 DD design, then apply the IV method. As shown in **Figure 3**, ride-hailing services were virtually non-existent in 2012, expanded rapidly around 2014, and had become behaviorally relevant in many cities by 2016, in all three measures we consider. Given this, we use the year 2012 and 2016 as the control and the treatment period, respectively. To test the null of no mediation effect ($AIE = 0$), we estimate the following TWFE regression on the two-year sample:

$$y_{cjtm} = \alpha_{ctm} + \sum_{\kappa \in \{c,pt\}} \beta_\kappa S_{cjt}(\kappa) + X'_{cjt}\gamma + \eta_{cjtm}. \tag{8}$$

As in eq. (1), y_{cjtm} is the monthly average (or maximum) NO_2 concentration in PUMA j of city c in year t and month m. Our baseline specification includes all city, year, month fixed effects; the full specification further includes the city-year fixed effect to directly control for the ride-hailing's market penetration status. We also include some time-varying covariates X_{cjt} that vary at the MSA-PUMA level. We discuss the validity of these controls as well as the results with varying sets of controls more fully in **Appendix I**. Our primary mediator of interest is the density-adjusted share of commuters $S_{cjt}(\kappa)$ using either private car ($\kappa = c$) or public transit ($\kappa = pt$):

$$S_{cjt}(\kappa) = s_{cjt}(\kappa)\frac{N_{cjt}}{A_{cj}},$$

where $s_{cjt}(\kappa) = (1/|\tilde{N}_{cjt}|)\sum_{i \in \tilde{N}_{cjt}} I_{it}(\kappa)$ is the share of commuters in the ACS sample $\tilde{N}_{cjt}$, $I_{it}(\kappa)$ is an indicator dummy for commuting mode κ, N_{cjt} is the number of commuters, and A_{cj} is the total land area in the PUMA j of city c in year t. Intuitively, S_{cjt} adjusts the raw commuter share s_{cjt} for the fact that there are more commuters per square mile of the area, and hence, there are more emissions coming from commuting per square mile of the area. Note that the share of commuters using other mode $S_{cjt}(o)$ must be dropped from eq. (8) since it is perfectly colinear with $S_{cjt}(c)$ and $S_{cjt}(pt)$ by definition.

In this regression framework, we test the null of no mediation effect by testing the null of $\beta_c = \beta_{pt} = 0$ against the alternative $\beta_c > 0$ and $\beta_{pt} < 0$. We also evaluate the null of the linear combination, $-\beta_c + \beta_{pt} = 0$, against the alternative hypothesis, $-\beta_c + \beta_{pt} < 0$: i.e., re-allocating the (density-adjusted) share of commuters from private car commuting to public transit commuting will decrease air pollution. The empirical challenge here is that S_{cjt} is not only endogenous but also likely correlated with ride-hailing's market penetration, the direct effect of which must be held constant. We thus need a set of instruments that would predict S, but is uncorrelated with both the unobservables η and the ride-hailing entry status D.

We construct our IVs in the spirit of Bartik shift-share instruments, exploiting Uber's regional rollout profile as the common *shift* and geography-based instruments as the city-specific *shares*. To measure Uber's regional rollout profile, we calculate the average exposure duration from Uber's official entry year to the treatment year (i.e., 2016) at the Census region level. To further strengthen identification, we leave out the city's own contribution. Specifically, we compute the leave-one-out regional exposure duration $\bar{\omega}^o_{r(c)}$ as:

$$\bar{\omega}^o_{r(c)} = \frac{1}{n_{r(c)} - 1} \sum_{i \neq c, i \in C_{r(c)}} (2016 - E_i^{\text{Uber}}),$$

where E_i^{Uber} denotes Uber's official entry year in city i, $C_{r(c)}$ the set of cities in city c's Census region r, and $n_{r(c)}$ is the number of cities in that region. Intuitively, $\bar{\omega}^o_{r(c)}$ captures the region's average exposure to ride-hailing rollout faced by city c, while excluding the city's own influence.

In this formulation, three features strengthen the plausibility of the exclusion restriction. First, the measure captures the regional rollout profile rather than city-specific rollout timing, making it less likely to be driven by city-specific local shocks. Second, we exclude each city's own contribution from the regional average, further reducing the influence of local shocks. Third, because the measure is constructed from Uber's official rollout timing rather than realized city-level penetration, the leave-one-out regional average is less mechanically tied to local exposure duration and can be interpreted as a counterfactual diffusion measure. Alternatively, one could construct a measure based on the regional average share of cities with Uber entry, again leaving out the city's own contribution: $\bar{\theta}^o_{r(c)t} = \frac{1}{n_{r(c)}-1} \sum_{i \neq c, i \in C_{r(c)}} D_{it}^{\text{Uber}}$, where D_{it}^{Uber} is an indicator for Uber's official presence in city i in year t. We prefer our average exposure duration measure, $\bar{\omega}^o_{r(c)}$, to the regional entry-share measure, $\bar{\theta}^o_{r(c)t}$, because the latter varies directly with contemporaneous Uber presence and may therefore be more closely correlated with city-level ride-hailing presence. Nonetheless, we report the results using $\bar{\theta}^{o}_{r(c)t}$ as well as other alternative instruments in **Appendix I**.

To arrive at our instruments Z_{cj}, we interact the leave-one-out regional exposure duration $\bar{\omega}^{o}_{r(c)}$ with geography-based instruments G_c (scaled by the size of each PUMA):

$$Z_{cj} = \bar{\omega}^{o}_{r(c)} \times \frac{G_c}{A_{cj}}.$$

We use the three geography-based instruments, the planned route IV, historical route IV, and inconsequential unit IV, which are known to predict city-level transportation networks, while remaining plausibly exogenous to contemporaneous local shocks conditional on the included covariates.

Why should Z_{cj} be valid instruments for the changes in the MSA-level commuting patterns? Note first that we need both cross-sectional and temporal variation in our instruments that would predict the intertemporal changes in commuting mode shares across cities. For cross-sectional variation, we use the geography-based instruments G_c. In the economic geography literature, these instruments are interpreted as capturing *counterfactual* transportation networks that would predict *actual* transportation infrastructure, while remaining plausibly exogenous to contemporaneous local shocks conditional on the included covariates (Redding and Turner, 2015). In our setting, they provide cross-city variation in transportation network structures that are relevant for commuting mode choice when ride-hailing diffuses locally. One could alternatively use pre-treatment commuting shares or public transit network density, which may appear more closely aligned with the spirit of Bartik instruments. However, those variables may also affect local air pollution directly, creating a more immediate exclusion concern. We therefore use these well-established geography-based instruments as our *share* component.

For temporal variation, we need a *shift* component that is closely correlated with the changes in commuting mode shares between 2012 and 2016, but is plausibly uncorrelated with city-level local shocks. Analogous to the geography-based IVs, our leave-one-out regional average exposure duration is designed to capture *counterfactual* exposure windows between 2012 and 2016 that would predict *actual* exposure duration during which city-level commuting patterns respond to ride-hailing services. On one hand, because regional exposure duration is closely correlated with city-level exposure duration, and ride-hailing penetration induces changes in commuting patterns, it is also correlated with the direction and magnitude of city-level commuting shifts. On the other hand, because our measure uses only Uber's official rollout timing, is aggregated at the Census-region level, and excludes the city's own contribution, it is less directly tied to realized city-level ride-hailing exposure duration and thus plausibly orthogonal to city-specific shocks, conditional on the included fixed effects and other time-varying covariates.

Interacting these share and shift components, therefore, generates the cross-sectional and temporal variation needed to predict changes in commuting mode shares across cities, while remaining plausibly orthogonal to unobserved city-level shocks affecting local air pollution. Intuitively, cities with similar exposure duration may respond differently depending on how conducive their transportation infrastructure is to substitution across commuting modes. The leave-one-out regional exposure-duration measure creates counterfactual variation in local exposure windows, while the geography-based instruments create counterfactual variation in transportation infrastructure. Together, these two components capture the relevant margins along which commuting patterns adjust to ride-hailing diffusion.

5.C. Instrument Validity for Mediation Analysis

We need to validate two conditions, exogeneity (or exclusion restriction) $cov(Z, \eta) = 0$ and relevance $cov(Z, S) \neq 0$, conditional on covariates. To see if the exclusion restriction is satisfied, we regress our outcomes on Z using only the untreated observations ($D_{2012} = D_{2016} = 0$), controlling for all covariates included in eq. (8). To confirm the relevance, we regress our mediator variables S on Z using all observations, again controlling for all covariates included in eq. (8). **Panel A of Table 3** reports the results of the former regressions while **Panel B** reports on the latter regressions. For each outcome, we run four regressions: the first three regressions use each IV individually as a regressor while the last regression uses all three IVs.

From **Panel A**, we see that our instruments are statistically insignificant particularly when used jointly. The F-statistic for the joint significance of all IVs is 1.87 and 1.27 (p-values: 0.178 and 0.320), respectively, for the monthly mean NO_2 and the monthly maximum NO_2. In the monthly mean NO_2 regression, some of our instruments are significant when used separately. We, however, believe that this is most likely due to a spurious correlation, for they turn highly insignificant once we use all three IVs. We thus conclude that our instruments are likely to satisfy the exclusion restriction. On the other hand, **Panel B** shows that all of our IVs are highly statistically significant both when used separately and when used jointly. The F-statistic for the joint test on all IVs is 22.48 and 171.45 (p-values: 0.000 and 0.000), respectively, for the (density-adjusted) share of commuters by private car and by public transit. There may be a concern that the first-stage results using the full sample may be driven by cross-strata rather than within-strata variation in our IVs. To address such a concern, we also report the first-stage results by density quartile in **Appendix I**. The results confirm that there is sufficient within-strata variation. In particular, the direction, magnitude, and statistical significance of IV coefficients are similar between the top quartile and the full sample. The formal test statistics for the weak IV and the overidentification are also reported in **Appendix I**. The Kleibergen-Paap rk Wald F-statistic for the weak IV diagnostic under clustered standard errors ranges from 9.678 to 14.961. For reference, the Cragg-Donald Wald F-statistic ranges from 797.012 to 906.091, well above the Stock-Yogo 10% maximal IV size critical value under homoskedasticity. These diagnostics indicate that our instruments have sufficient predictive power. Hansen's J-statistic for overidentification ranges from 0.079 to 0.483 (p-values range from 0.487 to 0.904), failing to reject the null of exogeneity. From these results, we conclude that our instruments plausibly satisfy all required assumptions for the causal mediation analysis.

5.D. The Mediating Effect of Commuting Modes on Air Quality

In this subsection, we report on the estimation results of two sets of regressions, which together provide support for the mediating mechanism on the air quality impact of ride-hailing. The first one is the OLS and the IV regressions based on eq. (8). As discussed above, this IV regression is intended to identify β_1 in the canonical equation system (6) — the average effect of commuting patterns on air quality, controlling for the direct and interaction effects of ride-hailing. The second set of regressions are intended to identify δ_2 in eq. (6) — the average effect of ride-hailing on commuting mode choice, controlling for all counfounds. In this set of regressions, we regress household-level choice indicators $I_{it}(\kappa)$ — which denote household i's commuting mode κ in year t — on *de facto* entry indicators D_{it}. To maintain consistency with eq. (8), we employ the same 2-by-2 DD design, restricting the sample to years 2012 and 2016 only. While this 2-by-2 DD design is simple and consistent with eq. (8), it is arguably heuristic and potentially subject to contamination bias. In **Appendix J**, we estimate the TWFE regressions using the full sample, leveraging the staggered DD design employed in the main outcome regression (1). We also report the results of unconditional event-study regressions, which show no evidence of differential pre-trends within the same population density cohort. Taken together, these two sets of regressions are intended to identify $\beta_1\delta_2$, the average indirect effect evaluated at $D = 0$, as in eq. (7). This helps us credibly establish the chain of causal relationships — from ride-hailing's market penetration to commuting mode changes, and then from commuting mode changes to air quality changes.

We first present the estimates of the effect of ride-hailing entry on commuting modes in **Figure 6**. Panels A, B, and C display the estimated effects on the probability of commuting by private car ($I(c)$), by public transit ($I(pt)$), and by other modes ($I(o)$), respectively. In each panel, we show the heterogeneous effect by MSA-level density quartiles as well as the overall impact (labeled "all"). For all regressions, we control for household-level covariates as well as MSA, PUMA, and year fixed effects. Standard errors are clustered at the PUMA level.

Our results suggest that on average, ride-hailing is estimated to decrease the share of private car commuters by 0.84 ppt, increase that of public transit commuters by 0.16 ppt and of other modes by 0.68 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). The effects are highly heterogeneous, however. In the highest density MSAs, the ride-hailing entry is estimated to decrease the share of commuting by private car by 1.11 ppt, increase that of public transit by 0.26 ppt, and increase that of other modes by 0.85 ppt. This implies that for the highest density MSAs, ride-hailing and public transit are complements to each other while they serve as a substitute for private car commuting. Furthermore, in the third-quartile MSAs, the estimates are significantly negative for private car commuting, insignificant for public transit commuting, and significantly positive for other modes. For the lowest and the second quartile MSAs, all estimates turn insignificant. These estimates are not only consistent with Hall *et al.*'s earlier findings, but also consistent with the pollution-decreasing effect of ride-hailing entry for these MSAs in **Table 2**.

We now turn to our main result of this subsection. **Figure 7** plots the OLS and IV estimates of $-\beta_c$ and $-\beta_c + \beta_{pt}$ from the TWFE regression (8). The estimate $-\hat{\beta}_c$ captures the average effect of quitting private car commuting while $-\hat{\beta}_c + \hat{\beta}_{pt}$ represents that of switching from private car to public transit. The OLS regressions control for the PUMA-level average household incomes, college education, share of Black and Hispanic population, and share of poverty share as well as MSA, year, and month fixed effects. For the IV regression, we use all three IVs in addition to controls. **Panel A** plots the estimates on the monthly mean NO_2 while **Panel B** plots using the monthly maximum NO_2. The standard errors are clustered at the MSA and PUMA level. In **Appendix I**, we provide a fuller discussion of the choice of these controls, report the full set of regression results under alternative sets of controls together with the corresponding first-stage statistics, and confirm that our IV estimates are robust to the alternative specifications.

The figure provides strong support for the mediating effect of commuting modes on air quality. The OLS estimates are both negative and statistically significant, with the magnitude of the estimate $-\hat{\beta}_c + \hat{\beta}_{pt}$ being greater than that of $-\hat{\beta}_c$. This implies that a higher commuter share of private car (public transit) is significantly associated with an increase (decrease) in NO_2 concentrations. Using our IVs, the estimates become even larger in magnitude and more statistically significant. This implies that the OLS estimates are biased toward zero. The magnitudes of the estimates are not economically small, either. To put it in context, our IV estimates suggest that if one percent of households were to switch from private car to public transit additionally, it would reduce the monthly average NO_2 concentrations by 0.30% in a community with average population density.

Lastly, we illustrate the counterfactual impact of ride-hailing's market penetration implied from our causal mediation analysis. In **Table 4**, we report the estimates of the implied indirect effect, $(\hat{\beta}_c\widehat{\Delta I(c)} + \hat{\beta}_{pt}\widehat{\Delta I(pt)})N/A$, for different levels of PUMA-level population density N/A, combining the estimates from **Figure 6** and **Figure 7**.[14] To reiterate, note that these estimates are intended to capture the estimate of AIE in (7). The top row shows the estimated effect of ride-hailing on raw commuting mode shares $\widehat{\Delta I(c)}$ and $\widehat{\Delta I(pt)}$ from **Figure 6** for each MSA-level population density quartile. In the left column, the mean, minimum, and maximum PUMA-level density N/A (those are rough estimates) are reported for each MSA-level density quartile. The matrix inside the table then shows the estimates of the implied mediation effect by applying the IV estimates from **Panel A of Figure 7** using these values in each column and row. For comparison, we also report the direct estimates from **Panel A of Table 2**. The estimates of the mediation effect range from +0.005 log point (in the lowest density MSA) to -0.048 (in the highest density MSA), and the magnitudes and signs of the estimates are similar to the direct estimates from **Table 2**. We see, however, important differences — when evaluated at the mean of the PUMA-level density, ride-hailing entry is estimated to reduce the monthly average NO_2 concentrations only by 0.003 log point in the high density quartile. This is less than 1/10 of the direct estimates reported in **Table 2**. The magnitude of the implied impact aligns more closely with the direct estimate when evaluated at the mean density of the top 30 PUMAs, and exceeds it when evaluated at the maximum PUMA-level density. **Table A16** in **Appendix K** reports alternative estimates of the implied average indirect effect under various assumptions. When evaluated at the mean density of the top 30 PUMAs, the estimates range from –0.020 to –0.030 — remarkably close to the direct estimate. While it is reasonable that the indirect effect is smaller than the total effect — since it excludes both the direct effect and the interaction effect of ride-hailing — we interpret the results as indicating either (i) that the direct estimates primarily come from the effects in high-density areas, or (ii) that additional behavioral pathways may exist through which ride-hailing directly influences ambient air pollution. We refrain from making any stance as to which is more likely. Nonetheless, we conclude from these results that the chain of causal relationships indeed exists such that ride-hailing's penetration into a highly dense city induces changes in workers' commuting modes, and the induced change in commuting modes in turn causally reduce transport-related air

[14]See **Appendix B** for an explanation of why this approach to calculating the mediating impact is conceptually consistent.

pollution there.

5.E. Other Pollution-reducing Mechanisms

Our results are consistent with our hypothesized mechanism outlined in **Appendix B** as well as previous studies that find (i) that ride-hailing tends to increases public transit use (Hall *et al.*, 2018) and (ii) that ride-hailing tends to reduce congestion and improve traffic speed in areas where the pre-treatment levels of public transit use are high (Krishnamurthy and Ngo, 2024). However, the estimates of the mediation effect (or the average indirect effect) reported here are small relative to the direct estimates reported in **Section 4**. Because the average direct effect and the interaction effect are not estimated in our mediation analysis, this may imply that there are other pathways through which ride-hailing reduces air pollution either directly without commuting mode shift or interactively with commuting mode shift (assuming our estimates are unbiased). Krishnamurthy and Ngo (2024) offer some findings that suggest the existence of such pathways. They find the congestion/pollution-relieving effect of ride-hailing mostly during off-peak hours and in less populated counties in California. This implies that ride-hailing also reduces air pollution through some other mechanisms than commuting mode change. For example, ride-hailing providers use newer, less polluting cars such as electric or hybrid vehicles. Furthermore, it is possible that the pollution-reducing effect of a shift in commuting mode is greater in the presence of ride-hailing services, indicating a potential interaction effect. For instance, commuters who shift away from private cars to public transit may also use ride-hailing for non-commuting purposes, potentially replacing more polluting vehicles such as older cars or buses. While such alternative pathways are worthy of further investigation, we do not have the data to fully explore such pathways, and thus, they are left for future research.[15]

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. A set of important questions arise from this debate:

[15] We also explored mechanism tests in the spirit of Kwon and Roth (2026), who develops a test of the sharp null of full mediation: The treatment affects the outcome only through the specified mediator or mediators. In our setting, however, applying their framework requires discretizing multiple continuous mediators and imposing additional assumptions. Because the resulting conclusions are mixed and highly sensitive to these choices, we treat the exercise as exploratory only and report the results in **Appendix L**.

(i) Does ride-hailing induce a behavioral change in commuters' mode choice?, (ii) Is the induced behavioral change large enough to affect air pollution in cities?, and (iii) Does this behavioral change tend to increase or decrease air pollution, in what kind of cities? We investigate these questions empirically by the following two-step approach.

Our first step is to credibly estimate the direct impact of ride-hailing entry on ambient air quality, exploiting staggered rollout of ride-hailing entry into U.S. cities. Though ours is not the first to investigate the question, we improve upon the earlier empirical design 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 employ stratification and imputation methods along with the heterogeneity-robust event-study estimator (Callaway and Sant'Anna, 2021; de Chaisemartin and D'Haultfoeulle, 2020). Our second step is to credibly estimate the average mediation effect of commuting mode change on ambient air pollution, building upon the literature on causal mediation analysis (Imai *et al.*, 2010; Frölich and Huber, 2017). To credibly identify the mediation effect, we construct a 2-by-2 treatment-control structure using an appropriately restricted subsample, and apply an instrumental variable method by employing a set of instruments that would predict commuting mode change, but not ride-hailing entry nor ambient air pollution, conditional on the included covariates.

Following these steps, our manuscript provides a set of complementary findings that reinforce each other. First, we find robust evidence that ride-hailing tends to reduce air pollution in large, dense cities, but has no significant effect in lower dense cities. Second, we also find robust evidence that ride-hailing tends to decrease private car use and increase use of public transit as a means to work, for these large cities. Third, we also find that the change in ambient air pollution is causally related to the change in commuting mode. We further demonstrate that the direct estimates of the pollution-reducing effect of ride-hailing in the first step are quite similar to the indirect estimates implied by the mediating effect of commuting mode. We emphasize here that to our knowledge, ours is the first to credibly establish this mediating mechanism empirically, although there is already a rich literature empirically investigating the environmental impact of ride-hailing per se.

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

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

Data Availability: The replication package for this article, including the data, code, and documentation, is available at Konishi and Ono (2026).

References

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

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

[3] Anderson, Michael L. and Lucas W. Davis (2026) Uber and Traffic Fatalities. *The Review of Economics and Statistics* 108 (2): 525–532. doi: https://doi.org/10.1162/rest_a_01385

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

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

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

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

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

[9] Callaway, Brantly and Pedro H.C. Sant'Anna. (2020) Difference-in-Differences with Multiple Time Periods. *Journal of Econometrics* 225 (2): 200-230.

[10] 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

[11] Castillo, Juan Camilo (2025) Who Benefits From Surge Pricing? *Econometrica* 93: 1811-1854.

[12] Chen, Stacey H., Yen-Chien Chen, and Jin-Tan Liu (2019) The Impact of Family Composition on Educational Achievement. *Journal of Human Resources* 54 (1): 122-170

[13] 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

[14] Chen, Y. T., Hsu, Y. C., & Wang, H. J. (2018) A Stochastic Frontier Model with Endogenous Treatment Status and Mediator. *Journal of Business & Economic Statistics* 38(2): 243–256.

[15] 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

[16] 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

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

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

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

[20] 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

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

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

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

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

[25] 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

[26] Fowlie, Meredith, Edward Rubin, and Reed Walker. (2019) Bringing Satellite-Based Air Quality Estimates Down to Earth. AEA Papers and Proceedings 109: 283-88

[27] 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

[28] Frölich, Markus, and Martin Huber (2017) Direct and indirect treatment effects–causal chains and mediation analysis with instrumental variables. *Journal of the Royal Statistical Society. Series B* (Statistical Methodology) 79(5): 1645–1666.

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

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

[31] Goodman-Bacon, Andrew. (2021) Difference-in-Differences with Variation in Treatment Timing. *Journal of Econometrics* 225 (2): 254-277.

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

[33] Grainger, Corbett, Andrew Schreiber, Wonjun Chang. (2019) Do regulators strategically avoid pollution hotspots when siting monitors? Evidence from remote sensing of air pollution. Revise and Resubmit at *American Economic Journal: Economic Policy*

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

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

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

[37] Heckman, James J., Hidehiko Ichimura, and Petra Todd. (1997) Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme. *Review of Economic Studies* Vol. 64, No. 4, pp. 605–654

[38] Huber, Martin (2019) A review of causal mediation analysis for assessing direct and indirect treatment effects. University of Freiburg/Fribourg Switzerland, Faculty of Economics and Social Sciences. *FSES Working Papers 500*

[39] Imai, Kosuke, Luke Keele, and Dustin Tingley (2010) A General Approach to Causal Mediation Analysis. *Psychological Methods* 15 (4): 309-334

[40] Imai, Kosuke, Luke Keele, and Teppei Yamamoto (2010) Identification, Inference and Sensitivity Analysis for Causal Mediation Effects. *Statistical Science* 25 (1): 51 - 71

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

[42] 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

[43] Konishi, Yoshifumi (2024) Economics of Ride-sharing. *Mitsubishi Economic Research Institute (MERI) Research Paper 157*, 111 pages (in Japanese)

[44] Konishi, Yoshifumi, and Ono, Akari (2026) Replication Package for: Is Ride-sharing Good for Environment? Evidence from Combining Satellite and Survey Data on U.S. Cities. *Journal of Urban Economics* [journal] Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], https://doi.org/10.3886/E246404V3

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

[46] Kwon, Soonwoo and Jonathan Roth (2026) Testing Mechanisms. *The Review of Economic Studies*, rdag028, https://doi.org/10.1093/restud/rdag028

[47] 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

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

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

[50] Powdthavee, Nattavudh, Warn N. Lekfuangfu, and Mark Wooden. The Marginal Income Effect of Education on Happiness: Estimating the Direct and Indirect Effects of Compulsory Schooling on Well-Being in Australia. *Melbourne Institute Working Paper No. 16/13*

[51] 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

[52] 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

[53] Rosaia, Nicola (2025) Competing Platforms and Transport Equilibrium. *Econometrica* 93 (6): 2235-2271. https://doi.org/10.3982/ECTA21773

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

[55] 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

[56] Sun, Liyang, Sarah Abraham (2020) Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects. *Journal of Econometrics* 225 (2): 175-199.

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

[58] 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

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

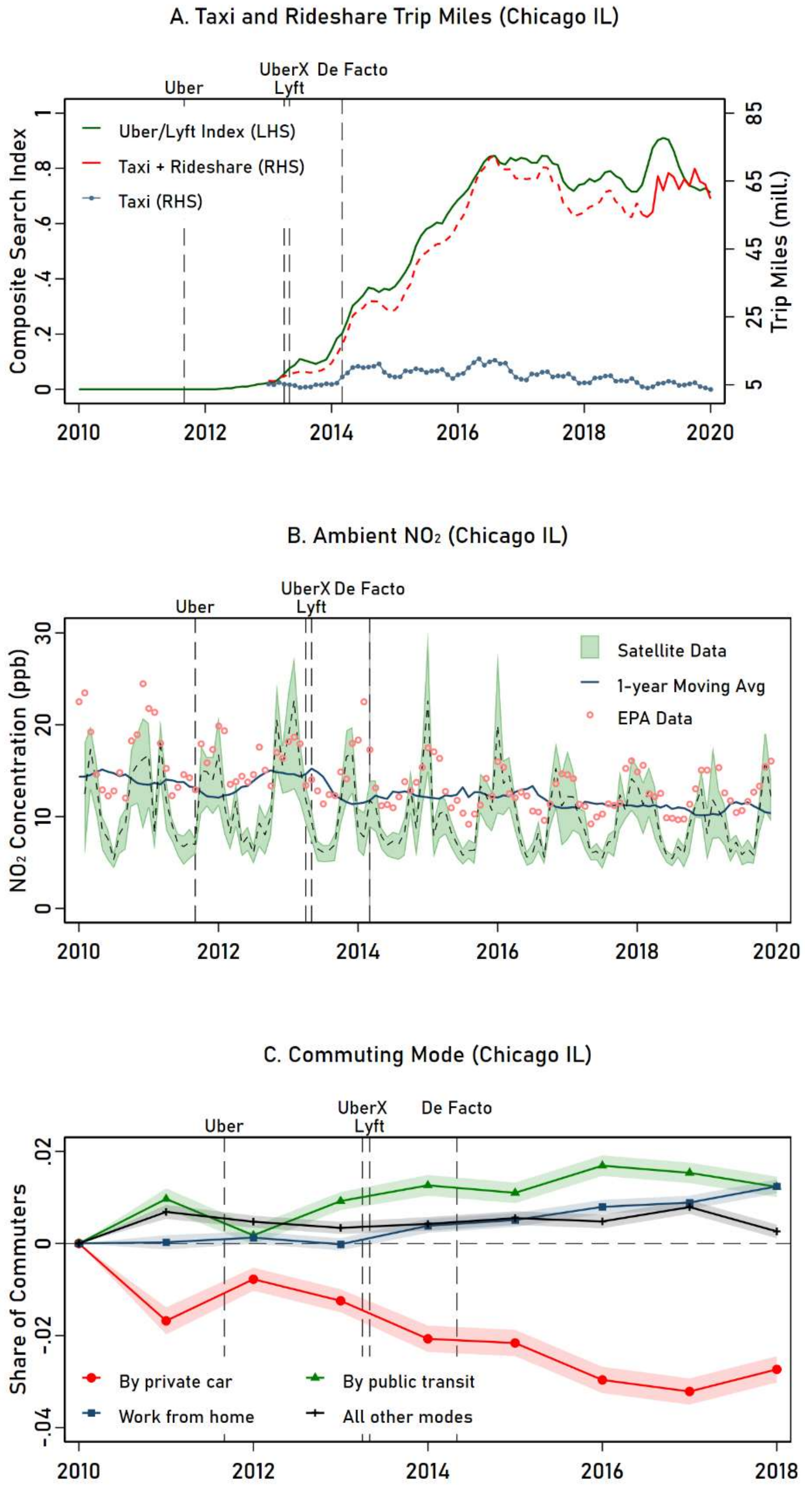
[60] VanderWeele, Tyler J. (2016) Mediation Analysis: A Practitioner's Guide. *Annual Review of Public Health* 37: 17-32

[61] 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

[62] Wooldridge, Jeffrey M. (2025) Two-way fixed effects, the two-way mundlak regression, and difference-in-differences estimators. *Empirical Economics* 69: 2545–2587.

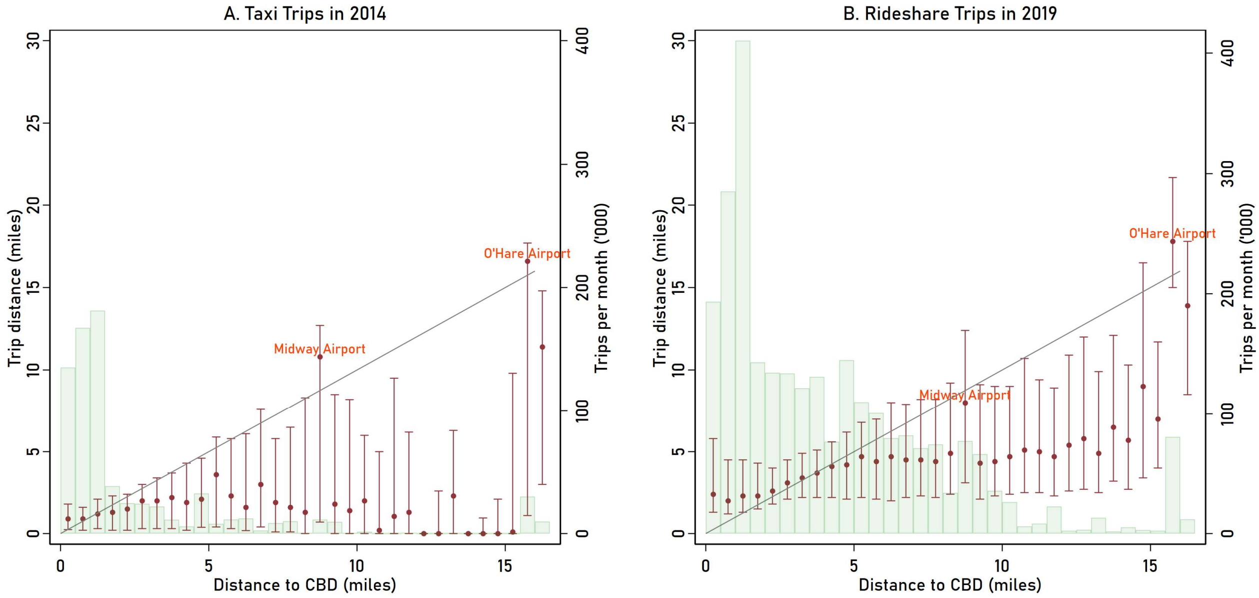
[63] Zou, Eric Yongchen (2021) Unwatched Pollution: The Effect of Intermittent Monitoring on Air Quality. *American Economic Review* 111 (7): 2101-26.

Figure 1. Ride-hailing Trips, Air Pollution, and Commuting Modes in Chicago, IL



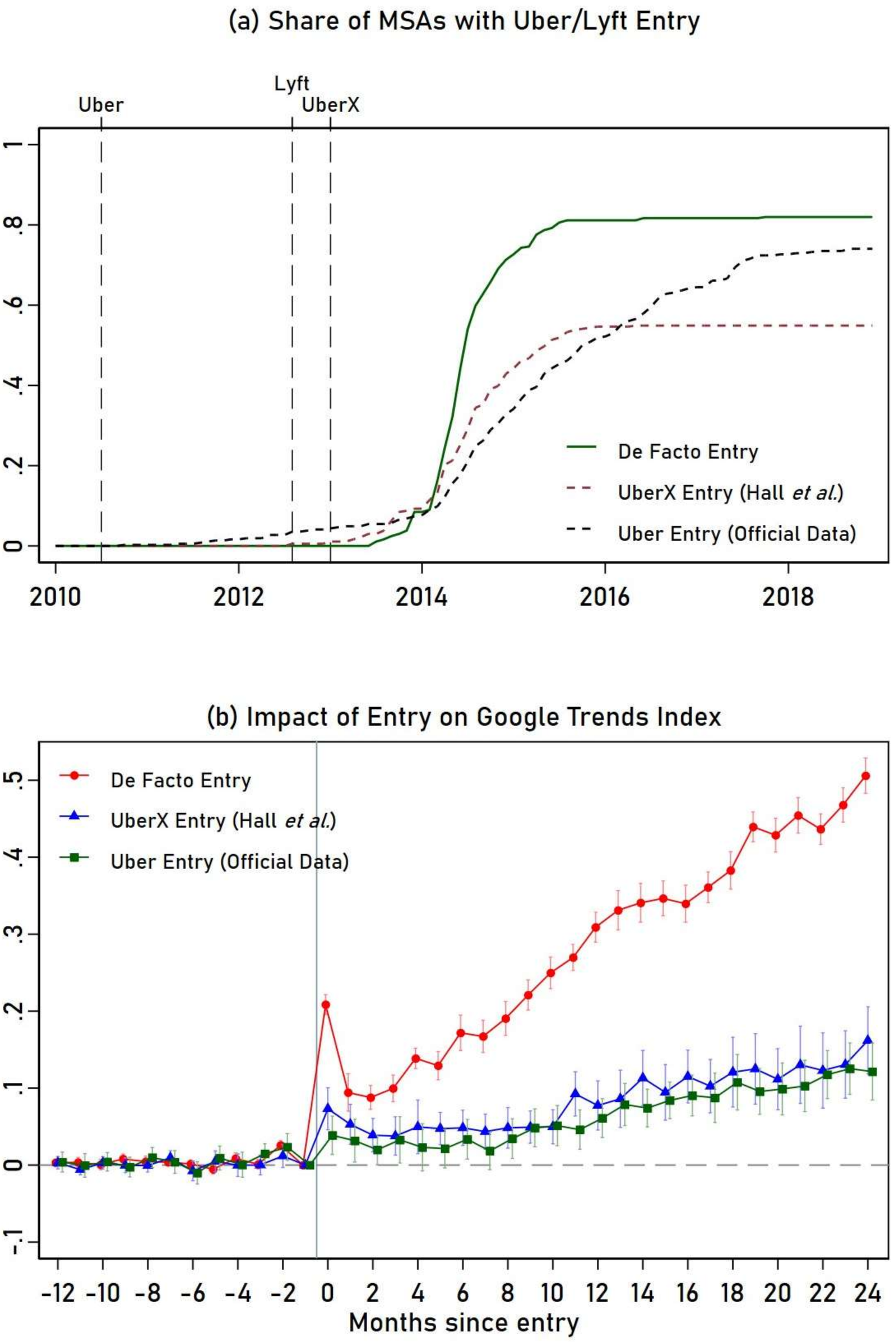
Note: Panel A plots the time trends of monthly taxi trip miles (blue), ride-hailing (taxi + rideshare) trip miles (red), Google Trends Index (green). The dashed line is the predicted ride-hailing trip miles obtained from a third-degree polynomial regression of rideshare trip miles on Google Trends Index. Panel B plots monthly MSA-level averages of ambient NO_2 concentrations using the satellite-based data (light green area) and EPA's monitoring data (red circles). Panel C plots the MSA-level commuter shares of alternative commuting modes to work relative to 2010 using the household-level data from the ACS. The details of the data are explained in Section 3.

Figure 2. Geographic Distribution of Ride-hailing Trips in Chicago, IL
Morning and Evening Rush Hours, Weekdays



Note: Panel A plots the average distance per trip of all taxi rides during the morning and evening rush hours against their pickup location bins measured in distance to the central business district (CBD) and the distribution of taxi trip volumes per month over the pickup locations, in 2014 (before ride-hailing's full market penetration) within the community area boundary of Chicago. Panel B plots the same for all TNC trips. The whiskers represent the 25th and the 75 percentiles of trip distance. The rush hours are defined as 6-9 am in the morning and 16-19 pm in the evening. These rush hours are determined from the observed peaks in the number of ride-hailing rides during the weekdays, and thus, may not necessarily match the rush hours in other transportation modes such as freeway traffic and public transit. The details of the data are explained in Section 3-F.

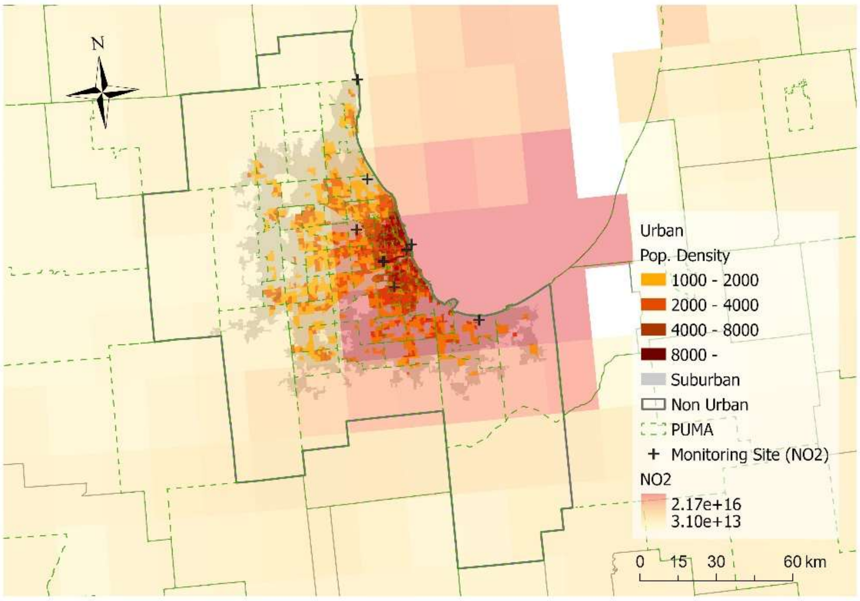
Figure 3. De Facto versus Reported Entry



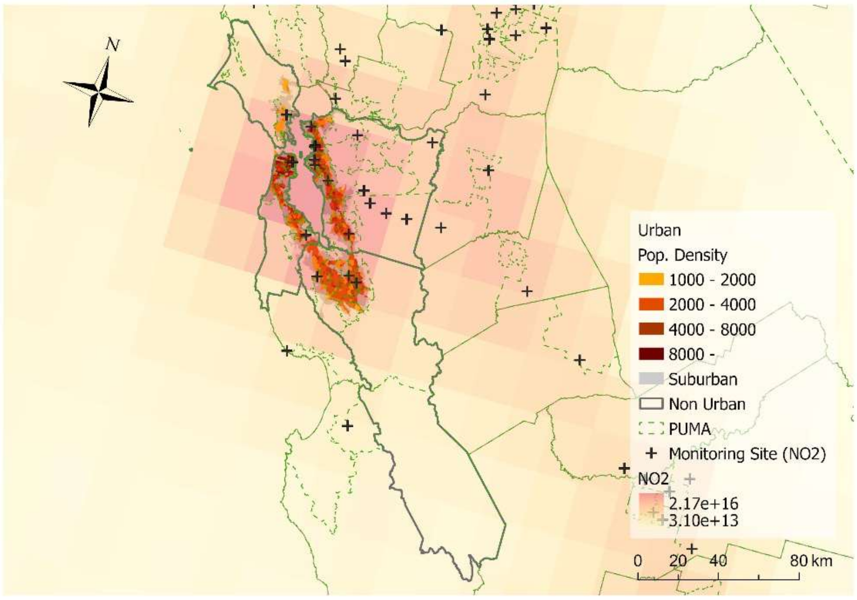
Note: The top panel report the cumulative distribution of ride-hailing entry using three alternative measures of entry: our measure based on Google Trends indices, UberX entry dates from Hall *et al.* (2018), and official Uber entry dates. The bottom panel report the event study estimates regressing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

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

A. Chicago, IL

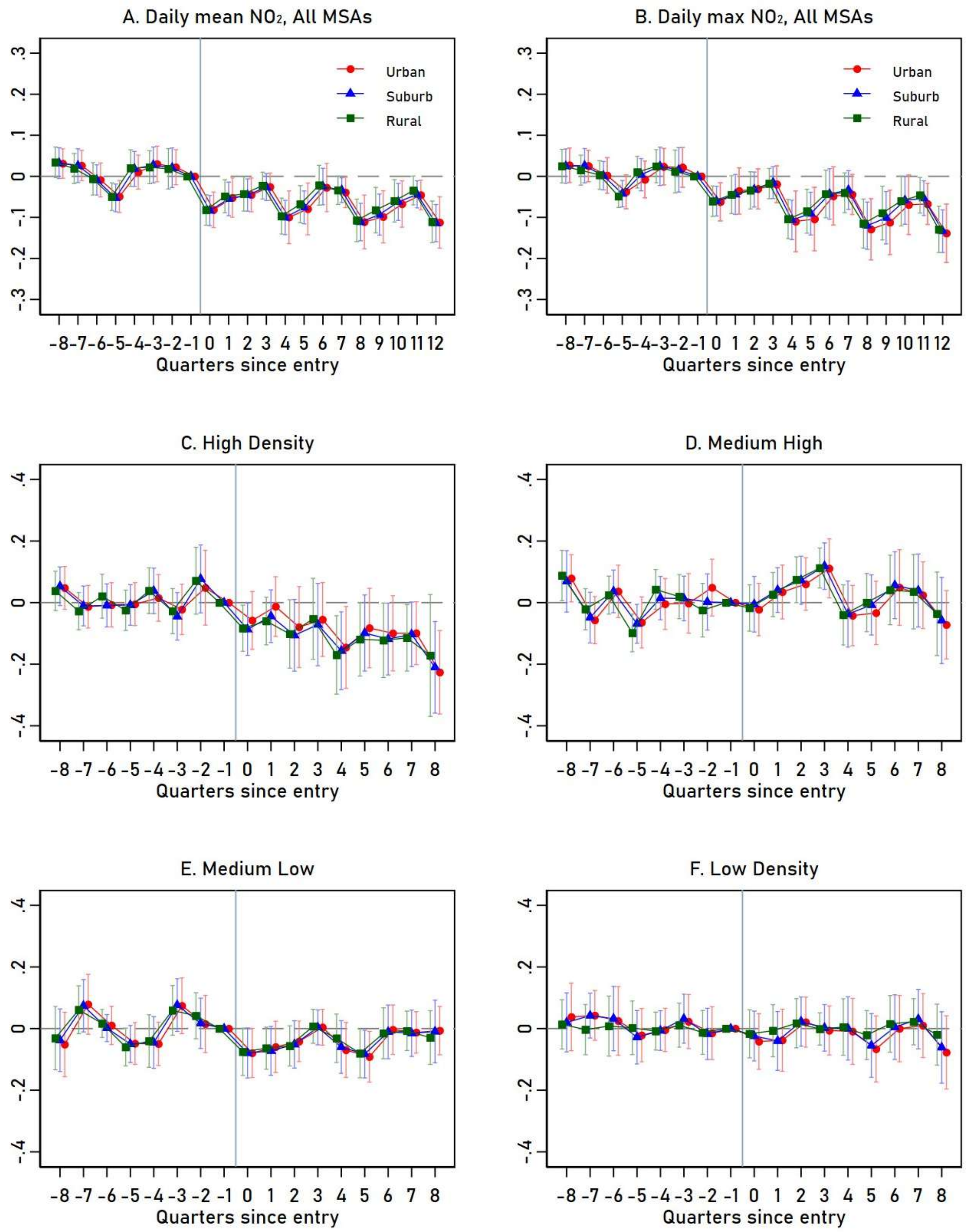


B. San Francisco, CA



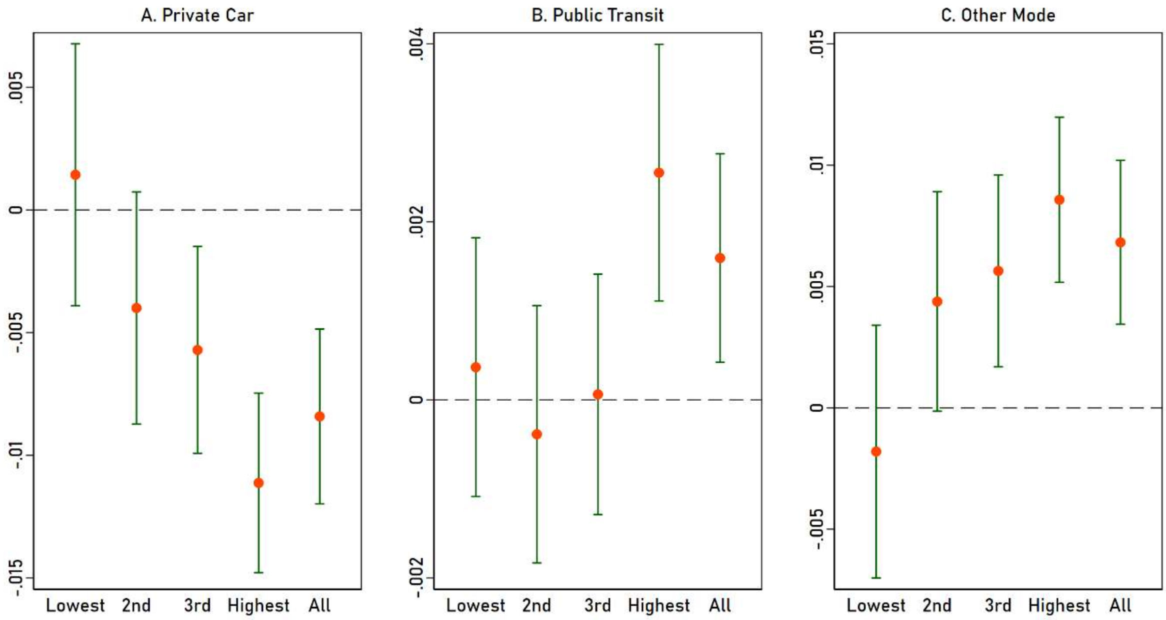
Note: The figures plot the grid-level monthly average NO_2 for Chicago, IL and San Francisco, CA in January, 2014, using data from the NASA's Goddard Earth Sciences Data and Information Services Center. The map also shows the MSA boundary (black and bold), the PUMA boundary (dashed green), and the geographic location of monitoring stations (+). For reference, the map also indicates the population density scale.

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



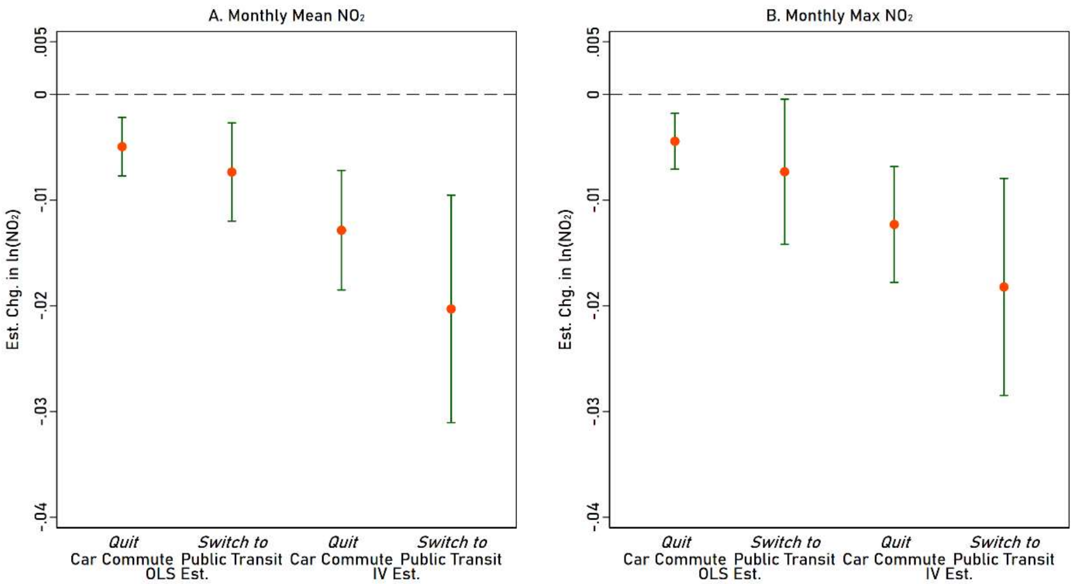
Note: The two graphs on the top panel report the event study estimates from the dCDH estimator on the pooled sample. The other four graphs plot the event study estimates on each population density quartile using monthly maximum NO_2 concentrations. Whiskers represent the 95% confidence intervals using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

Figure 6. Estimates of the Effect
of Ride-hailing on Commuting Mode Choice



Note: The figures plot the estimates of the effect of ride-hailing entry on commuting mode choice using household-level observations in the 2-by-2 difference-in-differences design. The whiskers represent 95% confidence intervals using robust standard errors clustered at the PUMA level. All regressions control for the household-level covariates such as household incomes, college education, racial and poverty dummies as well as MSA, PUMA, and year fixed effects.

Figure 7. Estimates of the Mediating Effect
of Commuting Modes on NO_2 Concentrations



Note: The figures plot the estimates of $-\beta_c$ and $-\beta_c + \beta_{pt}$ from the OLS and IV regressions in eq. (8). The whiskers represent 95% confidence intervals using robust standard errors clustered at the MSA and PUMA level. Panel A plots the estimates on the monthly mean NO_2 while Panel B plots using the monthly maximum NO_2. The OLS regressions use the PUMA-level average household incomes, college education, share of Black and Hispanic population, and share of poverty share as well as MSA, year, and month fixed effects. For the IV regression, we use all three IVs in addition to the controls.

Table 1. Pre-treatment Conditions by Treatment Status

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

Note: All statistics are as of year 2010.

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

A. Monthly Mean NO_2 Concentrations

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

B. Monthly Maximum NO_2 Concentrations

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

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

Table 3. Instrument Validity for Mediation Analysis

A. Exclusion Restriction

	Monthly Mean NO_2				Monthly Maximum NO_2			
L.O.O Regional Exposure ×Planned Route IV	108.882 ** (42.005)			48.052 (151.975)	54.848 (92.355)			280.095 (334.716)
L.O.O Regional Exposure ×Historical Route IV		155.289 ** (65.633)		109.42 (239.399)		45.65 (131.743)		-288.855 (503.264)
L.O.O Regional Exposure ×Inconsequential Unit IV			3.376 (2.333)	4.634 (2.960)			10.452 (5.983)	11.063 * (6.285)
F-statistic				1.87				1.27
Obs.	501	501	501	501	501	501	501	501

B. Instrument Relevance

	Commuting by Car				Commuting by Public Transit			
L.O.O Regional Exposure ×Planned Route IV	105.467 *** (40.129)			-909.051 *** (305.750)	154.732 *** (6.248)			95.437 * (50.661)
L.O.O Regional Exposure ×Historical Route IV		111.445 ** (43.769)		402.245 (289.204)		154.404 *** (5.773)		-66.887 (50.857)
L.O.O Regional Exposure ×Inconsequential Unit IV			140.311 ** (63.598)	633.729 *** (88.207)			157.426 *** (6.282)	129.505 *** (32.224)
F-statistic				22.481 ***				171.45 ***
Obs.	22,903	22,903	22,903	22,903	22,903	22,903	22,903	22,903

Note: Panel A reports on the results of placebo regressions, regressing PUMA-level monthly NO_2 concentrations on our instruments using only the untreated observations, controlling for PUMA-level averages of household-level covariates as well as MSA, year and month fixed effects. Panel B reports on the results of regressing PUMA-level commuting mode shares on our instruments using all observations and controlling for the same set of covariates. For all regressions, in parenthesis are robust standard errors clustered at the MSA-PUMA level.

Table 4. Evaluating the Mediating Effect of Commuting Mode Change

			Estimated Impact of Ride-hailing Penetration on Commuting Mode				
			MSA-level Population Density Quartile				
			Lowest	**2nd**	**3rd**	**Highest**	**Overall**
		ΔI_{car}	0.0014	-0.0040	-0.0057	-0.0111	-0.0084
		$\Delta I_{\text{pub. transit}}$	0.0004	-0.0004	0.0001	0.0026	0.0016
Population density							
MSA-level	**PUMA-level**						
Lowest	Mean	5.11	0.000	0.000	0.000	-0.001	-0.001
	Min	0.01	0.000	0.000	0.000	0.000	0.000
	Max	93.21	0.001	-0.005	-0.007	-0.015	-0.011
2nd	Mean	5.52	0.000	0.000	0.000	-0.001	-0.001
	Min	0.03	0.000	0.000	0.000	0.000	0.000
	Max	46.25	0.001	-0.002	-0.003	-0.007	-0.006
3rd	Mean	11.99	0.000	-0.001	-0.001	-0.002	-0.001
	Min	0.04	0.000	0.000	0.000	0.000	0.000
	Max	259.67	0.004	-0.013	-0.019	-0.042	-0.031
Highest	Mean	17.57	0.000	-0.001	-0.001	-0.003	-0.002
	Min	0.08	0.000	0.000	0.000	0.000	0.000
	Max	298.65	0.005	-0.014	-0.022	-0.048	-0.036
	Top 30	104.00	0.002	-0.005	-0.008	-0.017	-0.012
Direct estimates from Table 2			0.007	-0.010	-0.005	-0.034	

Note: The table compares the mediating effect of commuting mode change against the direct estimates of the effect of ride-hailing on monthly average NO_2 concentrations taken from Table 3. Each cell in the table reports the estimated impact $(\hat{\beta}_c \widehat{\Delta I}_c + \hat{\beta}_{pt} \widehat{\Delta I}_{pt}) N/A$ using estimates $\hat{\beta}_c$, $\hat{\beta}_{pt}$, $\widehat{\Delta I}_c$, and $\widehat{\Delta I}_{pt}$ from Figures 6 and 7 and evaluating at various PUMA-level density estimates.

Online Appendix for

Is Ride-sharing Good for Environment?
Evidence from Combining Satellite and Survey Data on U.S. Cities

By Yoshifumi Konishi and Akari Ono

This online appendix provides supporting material for the paper's interpretation, identification, and robustness analyses. It clarifies the treatment and outcome measures, reports additional diagnostics and specifications, and presents supplementary mechanism exercises. Appendix A reviews the related literature and clarifies our counterfactual. Appendix B develops the conceptual framework for how ride-hailing can affect emissions and congestion. Appendix C discusses the interpretation of the Google Trends-based measure and reports robustness to a refined entry definition. Appendix D validates the satellite-based NO_2 data and documents limitations of EPA monitoring data. Appendix E examines density and public transit as sources of heterogeneous treatment effects. Appendix F provides evidence on the orthogonality of de facto entry timing. Appendix G explains the geography-based instruments. Appendix H compares the binary treatment with a continuous-intensity measure. Appendix I reports full IV results, first-stage diagnostics, and alternative instruments. Appendix J estimates the effect of ride-hailing entry on commuting modes using the full sample. Appendix K reports implied indirect effects under alternative assumptions. Appendix L presents exploratory Kwon-Roth mechanism tests.

Table of Contents

Appendix A. Related Literature on the Environmental Concerns of Ride-hailing Service

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

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

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.

Another angle in which to dissect the related literature is the level of aggregation in data used. On one hand, studies based on disaggregate data tend to find that ride-hailing increases congestion in particular segments of transport demand (e.g., rush hours, urban freeways, populated counties) while decreasing congestion in other segments (e.g., non-rush ours, non-urban freeways, less populated counties) [Erhardt *et al.* (2019), Tarduno (2021), and Krishnamurthy and Ngo (2024)]. On the other hand, studies that rely on more aggregate data such as MSA-level or monthly-level data tend to find that ride-hailing increases use of mass transit (Hall *et al.*, 2018), decrease congestion (Li *et al.*, 2021), decrease ambient air pollution (Kim and Sarmiento, 2021), and vehicle emissions (Ward *et al.*, 2019). The key to reconciling these mixed findings is the heterogenous effects of ride-hailing on various segments of transport demand both within and across cities. That is, ride-hailing may decrease air pollution by encouraging a combined use of ride-hailing and mass transit in cities that have dense public transit networks, yet may increase congestion in some part of the cities, due to deadheading by ride providers or use of hired rides in place of mass transit for part of the daily trips (but not the entire sequence of daily trips). Our analysis is also intended to clarify this point.

Appendix B. Why Ride-hailing May Change the Aggregate Level and the Spatial Distribution of Transport-related Emissions?

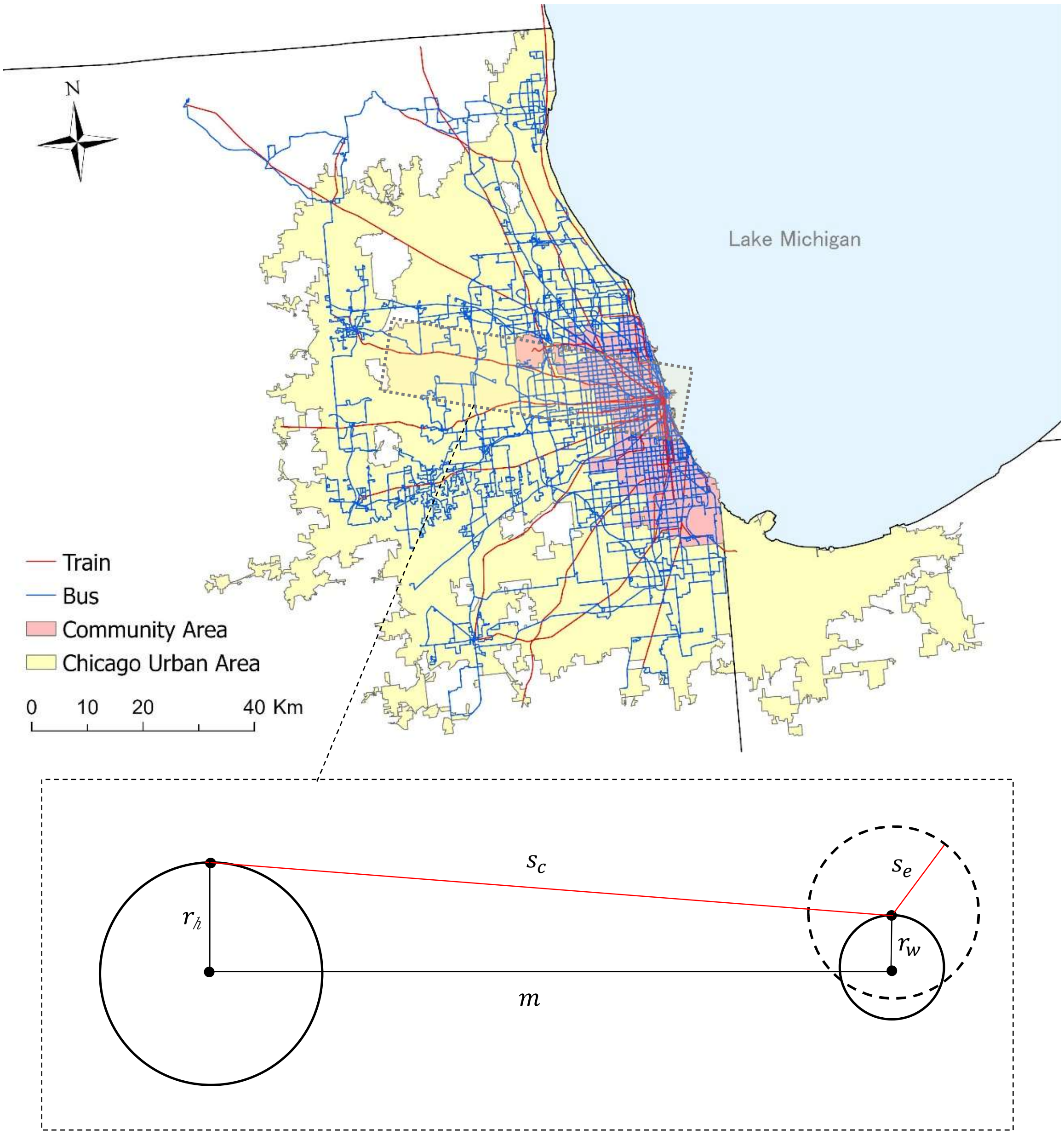
U.S. cities have experienced substantial suburbanization over the last half century. Central city population declined by 17% whereas total MSA-level population increased by 72% between 1950 and 1990 for large MSAs, due primarily to the rapid development of limited access highways over this period (Baum-Snow,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

This appendix demonstrates that ride-hailing can potentially reduce private driving more than simply replaces it while causing congestion around the city center using Chicago as an illustrative example of a monocentric city. **Figure A1** shows major train and bus routes in the urban and suburban areas of Chicago, along with a stylized diagram intended to conceptualize a certain segment of the city. As with many cities in the U.S., Chicago has a radial network of public transit routes that are highly concentrated around the central business district (CBD) and extend radially outward from the CBD to the suburban areas. As a result, many suburban areas have sparse transit networks that cannot directly connect their residents to either the CBD or other employment centers.

Imagine a commuter who lives in a suburb and commute to a city's employment center for work. Often, the nearest station is not located within a walking distance from/to commuter's home or workplace. Hence, the commuter may need to take another mass transit to the nearest station. Given this inconvenience, the commuter may opt to drive her car to work in the absence of low-cost, easy-to-hail ride services. This driving distance is denoted by s_c in the diagram. In addition, the commuter may use her car to run a few daily chores while at work: attending meetings at other places, buying things at stores, going to restaurants, etc. For simplicity, the driving distance serving such chores 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. Similar arguments apply to other kinds of commuting mode change that would reduce air pollution. For example, consider a commuter who lives near the CBD, but still sufficiently far from the CBD. The commuter could have walked or used bicycle (possibly in combination of public transit) to her workplace, but instead commutes by her own car mainly for the convenience of running daily chores at work. Ride-hailing services can induce such commuters to walking or bicycling (+ public transit) by allowing them to rely on hired rides instead of private cars for daily errands.

Figure A1. Complementarity Effect of Ride-hailing Entry on Daily Transport Choices



We have a reason to suspect that this kind of changes in commuting mode patterns might have occurred, at least in transit-friendly cities, although limited data are available to directly test such a hypothesis. **Figure A2** plots the hourly distribution of taxi trips by trip type in Chicago in 2014, the year in which ride-hailing has not yet fully penetrated the Chicago market whereas **Figure A3** plots the same for ride-hailing (TNC) trips in 2019, the year in which ride-hailing has fully penetrated the market. The figures demonstrate that while ride-hailing services increased trip volumes almost all hours on weekdays and weekends, they increased trip volumes much more during morning and evening peak hours on weekdays more than other hours/days, particularly outside CBD.

Figure A2. The Distribution of Taxi Trips by Hour of Day in Chicago in 2014

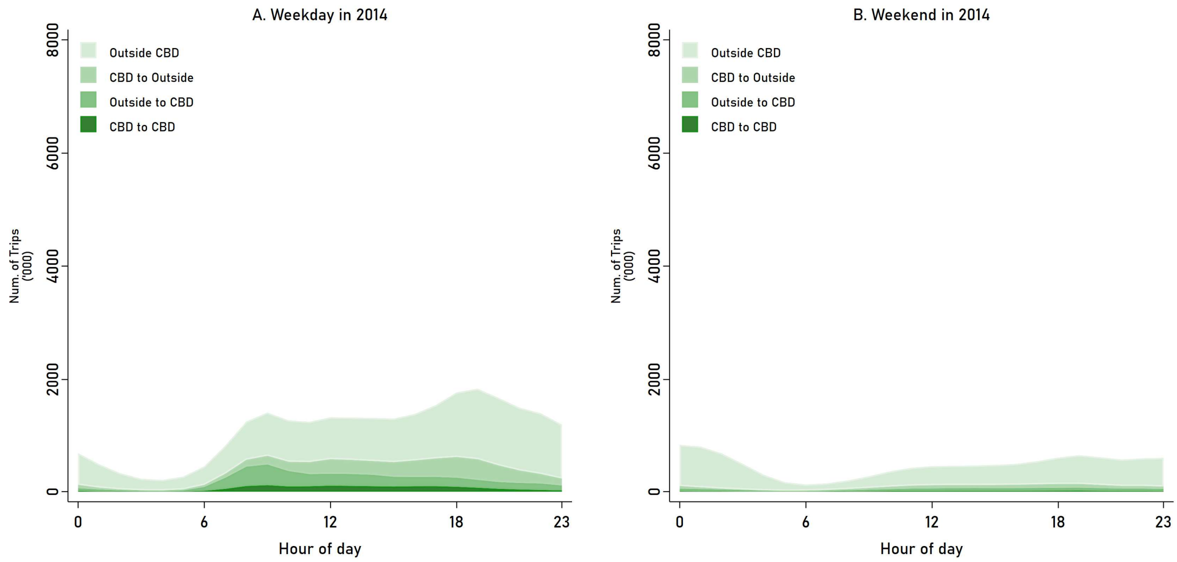
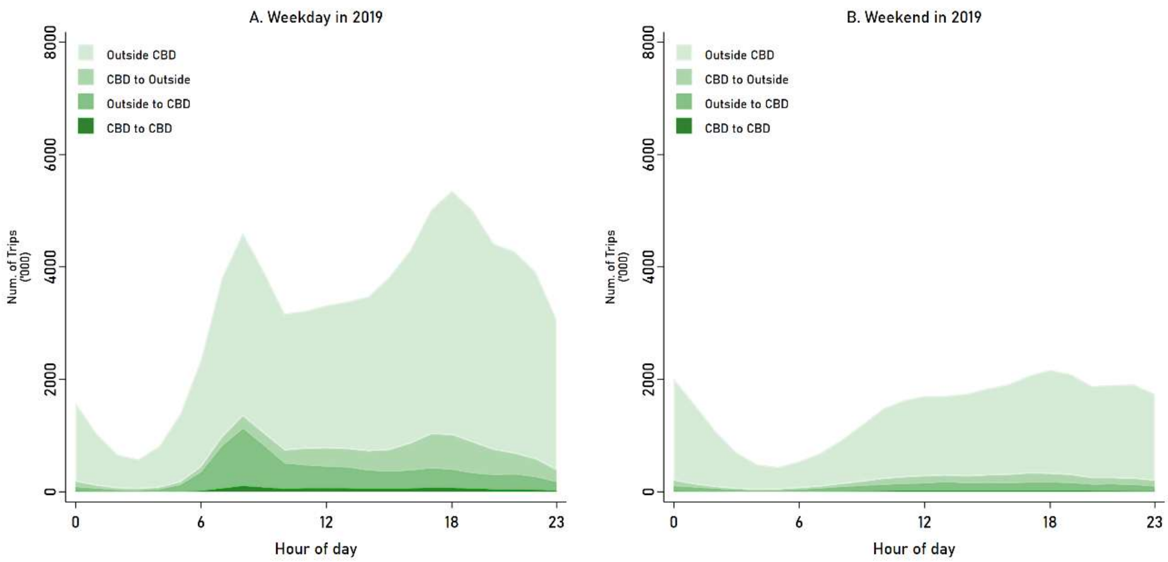


Figure A3. The Distribution of TNC Trips by Hour of Day in Chicago in 2019



When evaluating the impact of such city-wide modal shift, we need to account for another important aspect of ride-hailing services: There must be demanders as well as suppliers of the ride-hailing services. That is, upon entry of ride-hailing services, some commuters become *net demanders* while others become *net suppliers* of hired rides. In Appendix J, we show that ride-hailing entry increases the share of work-from-home commuters by 0.53 ppt and this is roughly half of the estimated decrease in the share of private car commuters (-1.26 ppt). These additional work-from-home commuters are presumably new ride-hailing providers.

To see why this is important in estimating the air pollution impact, consider the following simple example. Suppose that prior to ride-hailing entry, N identical commuters drive s_c+s_e, as in Figure A1. Suppose, further, that M of these commuters switch to the complementary use of ride-hailing and public transit (so their driving distance changes to r_h+r_w+s_e) and that the remaining $(N-M)$ commuters become the providers of ride-hailing services to these commuters, where the size of switchers M depends on matching technology/friction as well as the carrying capacity of hired rides. Assuming away deadheading, these ride-hailing drivers do not add driving distance and simply serve the driving distance r_h+r_w+s_e demanded by the M commuters. Thus, the air pollution impact of ride-hailing entry *per switched commuter* can be written as:

$$\begin{aligned}&\frac{M}{N}(r_h + r_w + s_e) - (s_c + s_e)\\ &\qquad = \frac{M}{N}(r_h + r_w - s_c) - \frac{N-M}{N}(s_c + s_e)\end{aligned}$$

That is, the pollution-reducing effect of ride-hailing consists of two terms: (i) the effect of replacing car commuting with combined use of ride-hailing with public transit and (ii) the effect of displacing private car use. Put differently, ride-hailing may reduce transport-related air pollution not only by encouraging public transit and other less-polluting modes, but also by simply inducing private car commuters to become ride-hailing providers.

Empirically, the pollution-reduction effect of ride-hailing ΔE can be estimated as:

$$\Delta E = \Delta I_{pub} dE/dI_{pub} + \Delta I_{car} dE/dI_{car}$$

where ΔI_{pub} and ΔI_{car} are the changes in the share of commuters using public transit and private car, respectively, and dE/dI_{pub} and dE/dI_{car} are the pollution impact of these commuting modes. This is the logic used in our mediation analysis in Section 5. Ideally, we obtain the estimates of all these terms by hour, by location (within city) and by city, exploiting exogenous variation in ride-hailing trip volumes. This can be done if we have (i) access to mobility data that are linked to the purpose as well as the hour, location, and city of the trips, and (ii) as good as random variations that would exogenously shift ride-hailing use. Unfortunately, neither seems feasible in our setting. We, therefore, opt for using the ACS data and estimate the average mediating effect of RH-induced commuting pattern changes exploiting variation at yearly and PUMA-MSA level, albeit all the well-known limitations of the ACS data. Furthermore, note that although the reasoning above assumes away deadheading or other effects of ride-hailing, such effects are directly controlled for by city-year fixed effect because they must be part of either the direct or the interaction effect of ride-hailing in eq. (6) of the main manuscript.

Appendix C. Interpretation of Google Trends Measure of Ride-hailing Presence

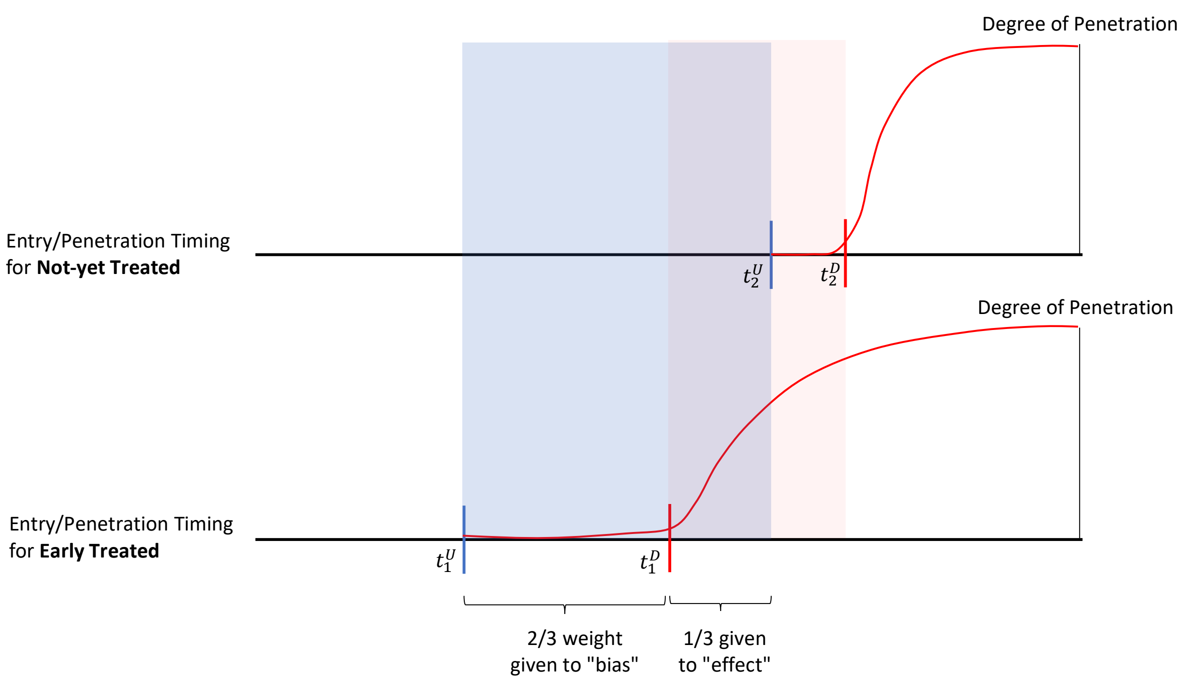
One of the novelties in our empirical study is the use of Google Trends indices to infer the time at which ride-hailing service becomes sufficiently salient and thus behaviorally relevant in the local transportation market. The official dates of entry may not accurately reflect that timing. For example, changes in transport choice would not have occurred on a meaningfully 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 becomes sufficiently salient in the city's market.

Our measure is designed to better align treatment timing in our DD design with the margin at which ride-hailing may begin to affect consumer behavior. In settings where search activity closely tracks actual local service usage, the measure may provide a better proxy for behaviorally relevant ride-hailing presence than official launch dates. At the same time, Google Trends may also reflect salience or anticipation in the absence of behaviorally relevant service availability. This appendix explains (a) the steps we have taken to minimize measurement bias, (b) the settings in which the measure appears strongest, (c) the extent of remaining bias in the data, and (d) the robustness of the main results to a more conservative entry definition.

Steps Taken to Minimize Bias: An important 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 positive' 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 exceed the pre-treatment maximum. Second, we hire a graduate student in 2020 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.

Pros and Cons of Google-Trends Measure: The most important merit of our measure is that it better aligns treated and untreated observations in our TWFE and event-study regressions, thereby reducing the bias in the estimated causal effect of ride-hailing on the outcomes of interest *if*, as we hope, it indeed accurately reflects actual local service usage. In the staggered DD design, these regressions are intended to compare the outcome trends of treated units against those of not-yet treated. Thus, if timings are incorrectly measured, these regressions will induce comparisons that are misaligned with the timing of behaviorally relevant ride-hailing presence. The resulting bias can be potentially severe in our settings because we expect the measurement errors to be heterogeneous and distributed across cities in a somewhat systematic way. It is known that Uber and Lyft penetrated different cities in different ways --- it took longer time to penetrate the cities entered early than the cities entered late because Uber and Lyft were already well known and also these companies had more know-hows by the time they entered into the later cities. To see how the potential bias arises more vividly, suppose the true penetration timings (t^D) lagged the official entry dates (t^U), and the lags are larger for cities entered early than cities entered late, as shown in **Figure A4** below.

Figure A4. ATT Estimates under Hypothetical Penetration Timings



Suppose that the degree of penetration follows the patterns shown by the red curves. The TWFE/event-study regressions using the official entry dates would estimate the estimated treatment effects by making comparisons between the early treated and the not-yet treated using the blue shaded period in the figure. In this hypothetical example, if correct ATT weights are used (such as in Callaway-Sant'Anna, 2021), then 2/3 of weights are given to "wrong comparisons". Assuming conditional parallel trends hold, 2/3 of weights are given to zero effects, so it tends to attenuate the true causal effect toward zero. If conditional parallel trends do not hold or hold partially, then 2/3 of weights are given to confounding bias. There are, of course, many other cases not discussed here (e.g., penetration timing can be even reversed), so in general, the magnitude and direction of bias is quite ambiguous. In contrast, the regressions using our *de facto* entry measure are intended to make comparisons in the red shared period.

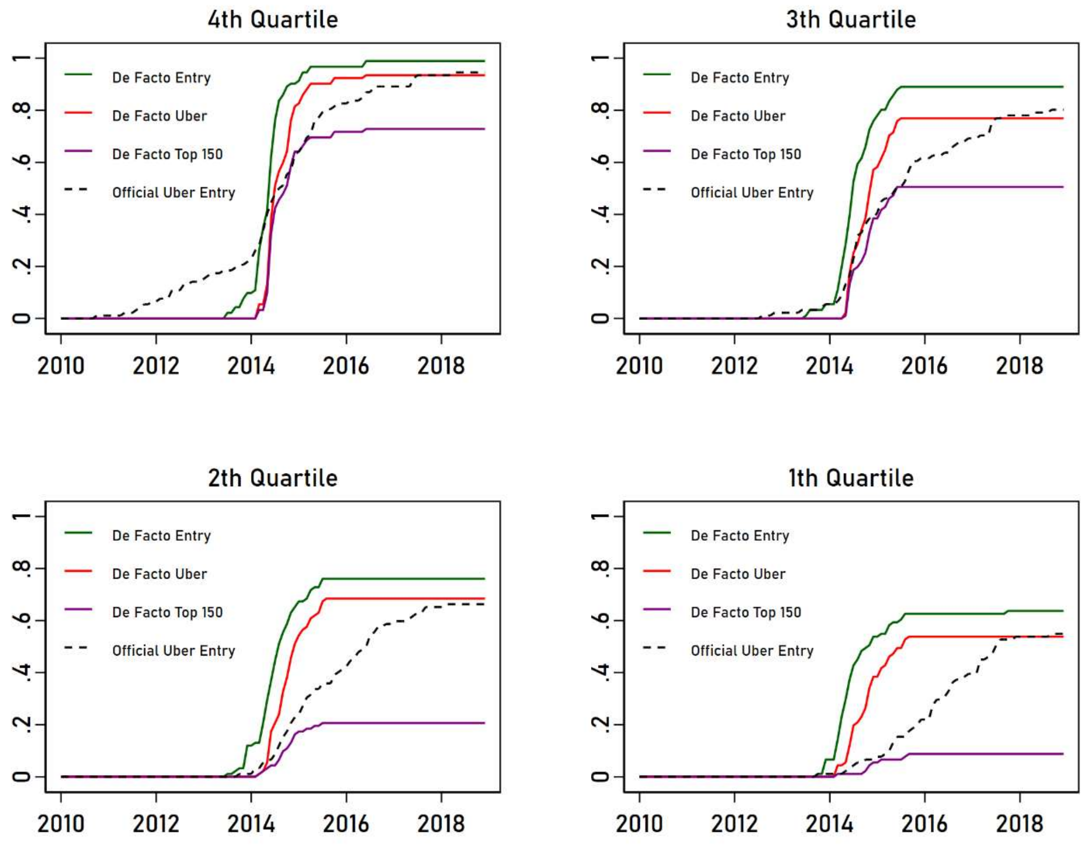
Another important merit of our entry measure is that it is likely exogenous conditional on covariates, unlike the other intensity measures such as ride-hailing trip volume, the number of ride-hailing drivers, or the Google Trend indices. It is widely recognized in the literature that these intensity measures are highly endogenous even after controlling for covariates because they are the equilibrium outcomes of individual decisions, and that this endogeneity is very hard to address because one would need an exogenous source of variation that would shift the *level of intensity of ride-hailing activity* across cities and over time, not just the timing and location of entry. On the other hand, the exact timing at which the Google Trends index crosses the threshold is difficult to predict conditional on observed MSA-level covariates and official entry timing. Thus, our *de facto* entry measure takes the merit of the Google Trends indices as intensity measures while alleviating their demerits.

An important limitation of our Google Trends-based measure is that it may not track actual local service availability or usage as closely as direct measures such as ride-hailing trip volumes or the number of active providers. As a result, the measure may not always identify the timing of behaviorally relevant ride-hailing presence with equal precision across cities. In particular, when search intensity reflects consumer interest or anticipation in advance of actual service availability, the measure may overstate the timing of entry. In fact, we can reverse the same argument in Figure A4: That is, if t^U is the *de facto* entry and t^D is the true penetration timing, then 2/3 of the weight is given to the bias! The concern is therefore not that the measure is uniformly invalid, but that its precision varies across settings.

Assessing the Extent of Salience and Anticipation: The discussion above suggests that we need to gauge the extent of such salience/anticipation in our data. While there is no definitive way to quantify the magnitude of such bias, an important concern arises when examining the distribution of entry timings shown in Figure 3(a) in the main text. By late 2015, approximately 80 percent of cities exhibit ride-hailing entry according to our *de facto* measure, compared with roughly 40 percent based on Uber's official launch data. Although our measure incorporates Lyft's entry in addition to Uber's, this discrepancy appears too large to be explained solely by the presence of Lyft or other smaller ride-hailing services.

To assess the extent to which Lyft's presence contributes to this discrepancy, we construct an alternative entry measure that excludes Lyft-related keyword searches and report the resulting distribution of entry timings by density quartile in Figure A5.

Figure A5. Cumulative Distribution of Entry Timings by Density Quartile



Conceptually, if the measure captures the timing at which ride-hailing becomes behaviorally relevant, it should not systematically precede official entry dates. When it does, the pattern may reflect salience or anticipation. Figure A5 shows that removing Lyft-related searches reduces the gap between our measure and official entry timing, particularly in the highest density quartiles, but does not eliminate it. This indicates that Lyft explains part, but not all, of the discrepancy. The remaining gap is concentrated in smaller, lower-density cities. Furthermore, a closer examination of Google Trends series at the city level reveals that these discrepancies are primarily driven by small cities exhibiting large fluctuations in search intensity over time. In such cases, Google Trends indices may reflect consumer interest, anticipation, or awareness rather than the actual usage of ride-hailing services.

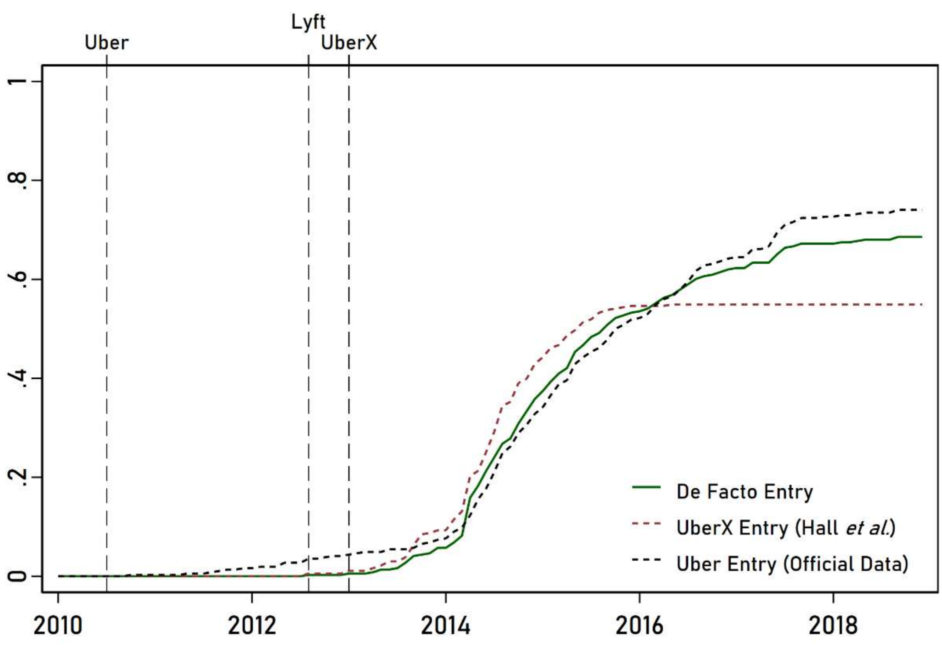
To explore this point further, the figure also reports the cumulative distribution of Google-based Uber entry counting only the largest 150 MSAs. Restricting attention to these largest cities substantially changes the picture: For these cities, the cumulative distribution of Google Trends-based entry lies below that of official Uber entry. This suggests that, in larger markets where actual Uber service already exists, the measure more closely tracks behaviorally relevant ride-hailing presence and may improve upon official launch dates as a proxy for the timing relevant to consumer response.

Estimation Results with A More Conservative Measure: The remaining concern is how robust our results are to a more conservative measure of entry. To further address this concern, we refine our *de facto* entry measure in the following manner:

$$Entry^R = \min(Entry^L, \max(Entry^D, Entry^U))$$

where $Entry^D$ is the entry date estimated using Google Trends data, $Entry^U$ is the entry date from Uber's official date, and $Entry^L$ is Lyft's entry date obtained from Hall *et al.* (2018). This refined measure gives us a conservative timing measure in the sense that *de facto* penetration always lags behind Uber's official entry in cities where Uber precedes Lyft. As shown in Figure A6 below, the refined entry measure closely traces out the official entry timings, but differs for cities entered early.

Figure A6. Cumulative Distribution of Entry Timings using Refined Entry Measure



We then estimate all the TWFE/event-study regressions using this *refined* entry measure. To make the comparisons easy, we report the results in a manner analogous to those of the main text. Table A1 reports the TWFE regressions using the refined entry measure. As shown, the estimates are quite similar to those reported in Table 2 in the main text. The estimates are negative, statistically significant on the highest density quartile and mostly insignificant for 1st – 3rd quartiles, for both monthly NO_2 means and maximums. The magnitudes are similar on monthly NO_2 means, but slightly smaller on monthly NO_2 maximums. Turning to the events-study estimates in Figure A7 (we use monthly NO_2 maximums as in Figure 5 in the main text), we again observe the same pattens: Overall, air pollution declines upon ride-hailing entry and magnitude gets larger over time, but we only see the declining trend in the highest density quartile. Thus, our results seem robust even accounting for measurement errors in our *de facto* entry measure.

Taken together, the evidence suggests that our Google Trends-based measure is most informative in large, early-entry cities, where search activity appears to track behaviorally relevant ride-hailing presence relatively closely and may do so more accurately than official launch dates. In smaller or later-entry cities, however, the measure may also reflect salience or anticipation in the absence of comparable service availability. Accordingly, estimates based on this measure should be interpreted more cautiously for those cities.

Table A1. TWFE Estimates of Effect of Ride-hailing Entry on Ambient Air Quality using Refined Entry Measure

A. Monthly Mean NO_2 Concentrations

	Mean NO_2 Concentration								
	Urban Area			Suburban Area			Non-urban Area		
Entry									
× 1st QuartileMSAs	-0.011	-0.006	-0.008	-0.008	-0.003	-0.005	0.013	0.019*	0.016
	(0.014)	(0.014)	(0.014)	(0.014)	(0.013)	(0.014)	(0.010)	(0.010)	(0.010)
× 2nd Quartile MSAs	-0.02	-0.016	-0.016	-0.021	-0.016	-0.016	-0.018	-0.013	-0.012
	(0.012)	(0.011)	(0.011)	(0.012)	(0.011)	(0.011)	(0.011)	(0.010)	(0.010)
× 3rd Quartile MSAs	-0.017	-0.013	-0.012	-0.018	-0.015	-0.013	-0.019 *	-0.016	-0.014
	(0.011)	(0.011)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
× 4th Quartile MSAs	-0.044 ***	-0.039 ***	-0.037 ***	-0.044 ***	-0.039 ***	-0.037 ***	-0.045 ***	-0.041 ***	-0.039 ***
	(0.011)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.009)	(0.009)
Controls included:									
Year/Month FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
MSA FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Climate controls		✓	✓		✓	✓		✓	✓
Other controls			✓			✓			✓
Obs.	38,911	35,023	35,023	39,309	35,421	35,421	39,254	35,367	35,367
Adj. R^2	0.759	0.757	0.757	0.765	0.762	0.762	0.780	0.778	0.778

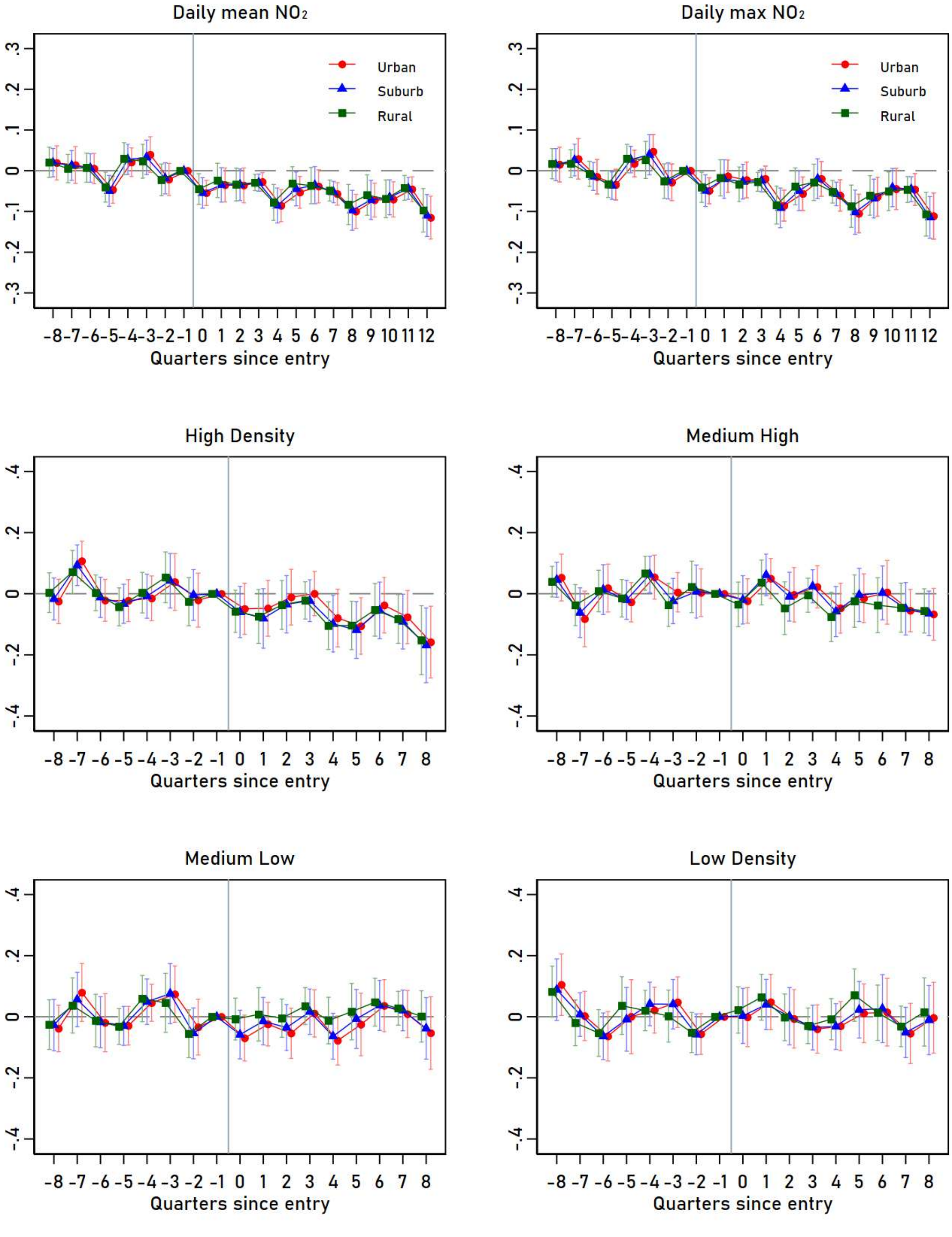
Note : Clustered robust standard errors in parenthesis.

B. Monthly Maximum NO_2 Concentrations

	Max NO_2 Concentration								
	Urban Area			Suburban Area			Non-urban Area		
Entry									
× 1st QuartileMSAs	-0.005	0.001	-0.002	-0.004	0.001	-0.001	0.014	0.019	0.016
	(0.018)	(0.019)	(0.018)	(0.018)	(0.019)	(0.018)	(0.013)	(0.014)	(0.013)
× 2nd Quartile MSAs	0.001	0.007	0.007	-0.005	0.000	0.001	-0.006	-0.002	-0.001
	(0.015)	(0.014)	(0.014)	(0.015)	(0.015)	(0.014)	(0.013)	(0.012)	(0.011)
× 3rd Quartile MSAs	-0.015	-0.01	-0.008	-0.020* *	-0.015	-0.012	-0.024 **	-0.019 *	-0.017 *
	(0.010)	(0.011)	(0.010)	(0.010)	(0.010)	(0.009)	(0.009)	(0.010)	(0.008)
× 4th Quartile MSAs	-0.042 **	-0.035 *	-0.033 *	-0.044 **	-0.038 **	-0.036 **	-0.047 ***	-0.043 ***	-0.041 ***
	(0.016)	(0.016)	(0.015)	(0.015)	(0.014)	(0.014)	(0.013)	(0.012)	(0.012)
Controls included:									
Year/Month FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
MSA FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Climate controls		✓	✓		✓	✓		✓	✓
Other controls			✓			✓			✓
Obs.	38,911	35,023	35,023	39,309	35,421	35,421	39,254	35,367	35,367
Adj. R^2	0.662	0.660	0.660	0.679	0.677	0.677	0.712	0.709	0.709

Note : Clustered robust standard errors in parenthesis.

Figure A7. Event Study Estimates of
the Effect of Entry on NO_2 Concentrations using Refined Entry Measure

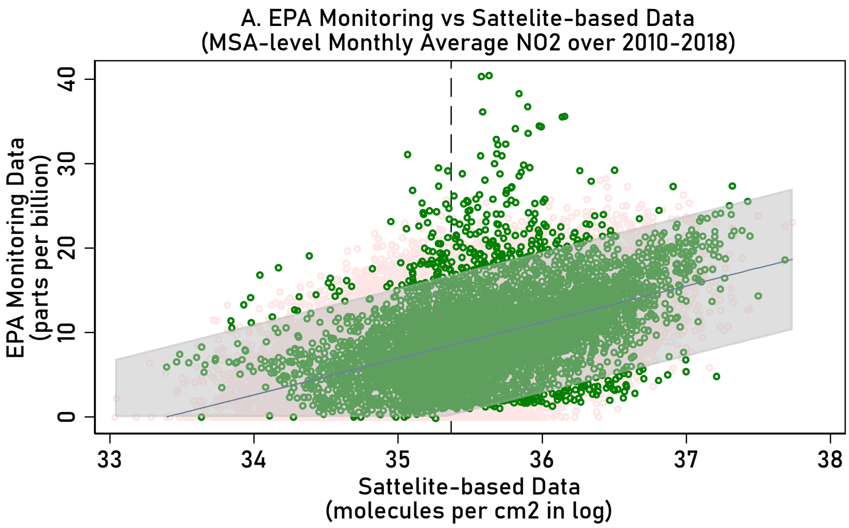


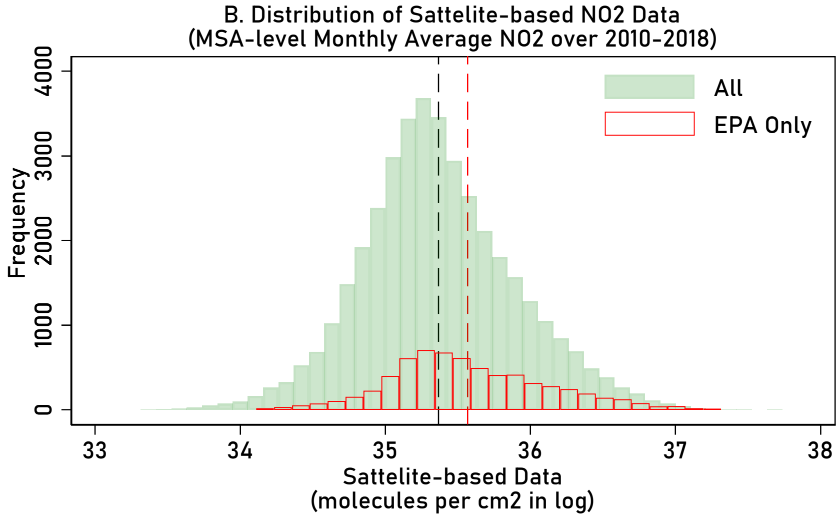
Appendix D. Satellite-based versus EPA Monitoring Data

As discussed in the main text, our satellite-based air pollution data allow us to cover a larger number of MSAs consistently over a longer time horizon --- the EPA monitoring data only cover 16.8% of our sample. Two important questions arise: (1) Are the satellite-based data likely a valid substitute when and where the EPA monitoring data are unavailable? (2) Is there an indication of bias such that EPA data might be missing at non-random, as argued in Zou (2021) and Grainger (2021).

Figure A8 examines the relationship between EPA monitoring data versus NASA's satellite-based data and the extent of the bias in the EPA monitoring data. Panel A displays a scatter plot of EPA monitoring data against satellite-based NO_2 data. Green dots use the observed data for which EPA data are available (i.e., 16.8% of our sample). Light red dots use the simulated values from a linear regression (blue line) in order to fill the missing EPA monitoring values. Panel B plots a histogram of satellite-based NO_2 data using the full sample (green) and the subsample for which EPA data are non-missing (red). The dashed lines represent the sample means. The figures show that there is a clear (statistically significant) relationship between the two, yet there is also substantial variation in the EPA's NO_2 data for each level of satellite-based NO_2 concentration. Unlike in Zou (2021) and Grainger (2021), however, there is an indication that EPA's NO_2 concentration data are slightly skewed toward higher concentration episodes. This is because the satellite data cover larger areas while the EPA monitoring stations are located around the city centers, so that the EPA monitoring data tend to miss low concentration episodes in our settings.

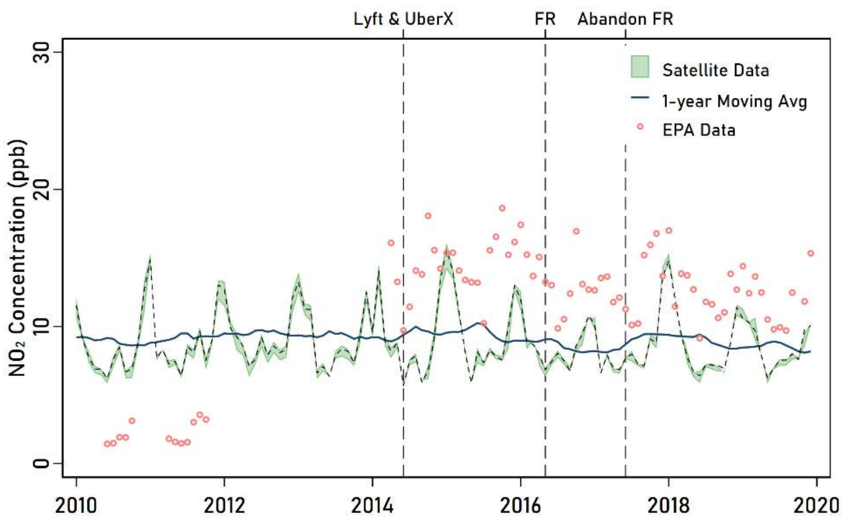
Figure A8. Satellite-based versus EPA Monitoring Data





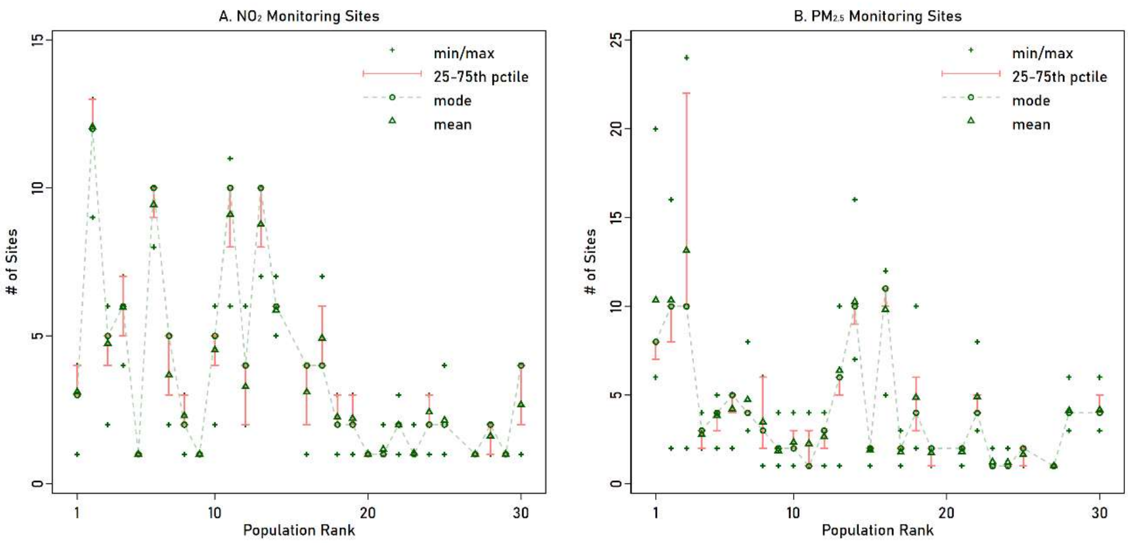
The extent of overall bias presented in **Figure A8**, however, understates the severity of the bias in EPA monitoring data. To see this, see **Figure A9**, which plots the time series of monthly MSA-level averages of ambient NO_2 concentrations in Austin, TX, using the satellite-based data (light green area) and EPA monitoring data (red circles). The figure plots the area-wide averages for non-urban, suburban, and urban areas in a manner similar to **Panel B of Figure 1** in the main text. The dashed line indicates the suburban NO_2 concentration levels, with the light green band bracketed by the urban and the non-urban NO_2 trends. The figure demonstrates a failure in statistical inference if we are to use EPA monitoring data. There are only two active NO_2 monitoring sites in Austin during the study period. With some intermittence, EPA shifted the location of the monitoring sites to a more urbanized area in 2014. This caused a huge jump in average NO_2 concentrations. However, looking at the satellite-based data, we don't see any sign of such a dramatic jump in NO_2 concentrations. This kind of bias in EPA data is likely to severely contaminate our estimates.

Figure A9. Air Quality Trends in Austin, TX: EPA vs. Satellite Data



There are still other questions that remain. First, previous studies use EPA monitoring data, mostly on larger cities. Typically, more monitoring sites are available in these large cities. So the question is, can we use the EPA monitoring data on NO_2 and other pollutants at least on large cities for credible inference and for robustness checks? The answer seems No. We have checked the EPA monitoring data on NO_2 and $PM_{2.5}$ for 30 largest MSAs (in terms of population size as of 2010) individually. We see that even in such large cities, (i) the number of monitoring sites differs substantially across cities, (ii) it increases and decreases rather frequently in some systematic way such that attrition is plausibly non-random, (iii) in some cities, there are months in which no data are reported, (iv) the changes in the number of monitoring sites are associated with the changes in the average NO_2 trends, and (v) we observe (i)-(iv) for both NO_2 and $PM_{2.5}$, but the attrition patterns are quite different between NO_2 and $PM_{2.5}$. The figures showing these patterns are available upon request. To concisely visualize the nature of the problem, **Figure A10** plots the distribution (mean, mode, 25-75-th percentiles as well as minimum and maximum) of the number of monitoring sites by population size ranking for the 30 largest MSAs over the study period. As shown, the number of monitoring sites differ substantially across MSAs, is not correlated with the size of cities, and varies frequently over time for each city. We also see there is no systematic pattern that can explain why some cities have changed the number of monitoring sites.

Figure A10. Distribution of EPA Monitoring Sites by Population Size,
30 Largest MSAs, 2010-2018



Second, the satellite-based air quality data are estimates based on satellite images. They may not be as precise as monitoring data if such monitoring data were available on all locations and times. So the question is, do we have sufficient variation in our satellite-based NO_2 data for credible inference, particularly on pollution peaks (i.e., monthly maximum) and on small-sized cities? The answer seems Yes. To verify this point, we plot the time series of monthly means and monthly maximums from the satellite-based NO_2 data (fitted estimates) for 30 largest MSAs and 30 smallest MSAs in **Figure A11** and the same from the EPA data in **Figure A12.** We see that on the satellite-based data, (i) monthly maximums are much higher on average and have larger variance than monthly means, (ii) pollution levels are generally higher and have larger variance in large MSAs than in small MSAs. We have sparse observations in EPA data, particularly on small MSAs. Nonetheless, comparing the two data, we see that the cyclical trends, the levels, and the variability of both monthly means and maximums seem quite similar between the satellite-based and the EPA data.

Figure A11. Time Series in Satellite-based Air Quality Data
in 30 Largest MSAs and 30 Smallest MSAs

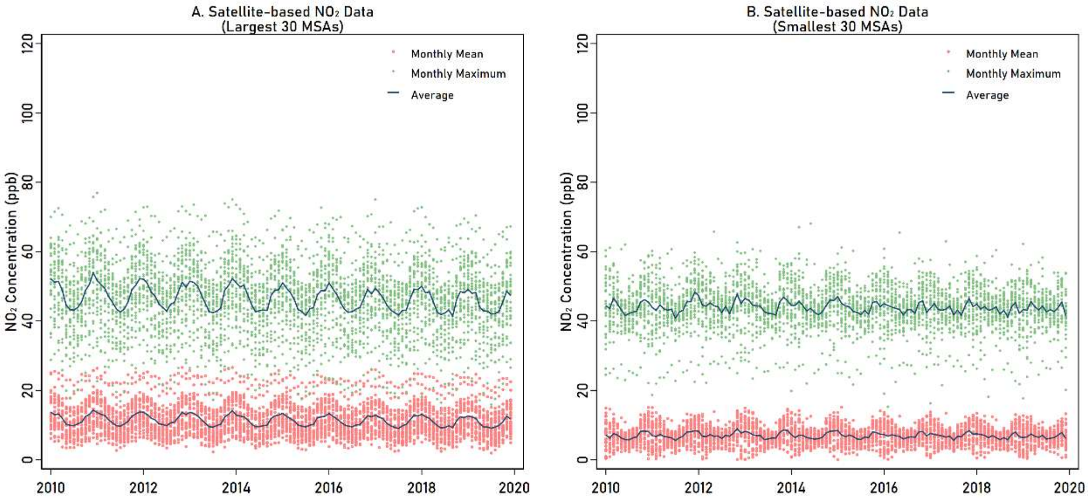
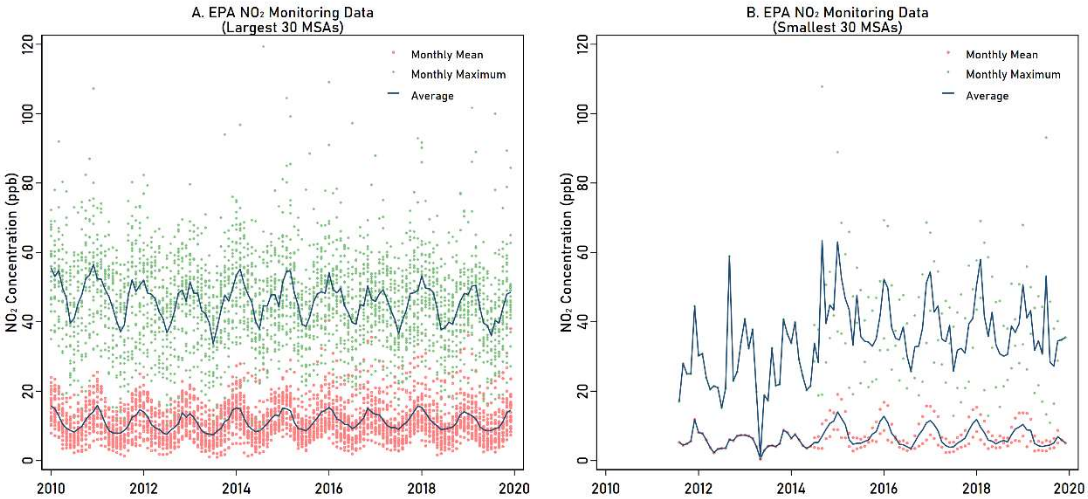


Figure A12. Time Series in EPA Air Quality Data
in 30 Largest MSAs and 30 Smallest MSAs



Because the EPA data seem to have the non-random attrition problem, we cannot conduct the same analysis using the EPA data for robustness checks even for the large cities. For example, the dCDH/CS event-study estimators fail to return estimates for most of the post-treatment periods even for the large cities. This also implies that even though we can still run the TWFE regression, the regression would not be making right comparisons. Nonetheless, in Table A2 and Table A3 below, we also report the results of TWFE regression of ride-hailing entry on air quality using EPA's monitoring data on NO_2 and $PM_{2.5}$ on two subsamples (the highest population density quartile MSAs and 30 largest MSAs). The signs and magnitudes of the estimates are similar to our main results, although we refrain from making any conclusion from the results for all the reasons explained above.

Table A2. TWFE Regression of Ride-hailing Entry on Air Quality using EPA NO_2 Data on High Density/Large MSAs

	4th Density Quartile MSAs					
	NO_2 Monthly Mean			**NO_2 Monthly Maximum**		
Entry	-0.032 (0.030)	-0.033 (0.030)	-0.044 (0.030)	-0.027 (0.019)	-0.027 (0.019)	-0.025 (0.021)
Controls included:						
Year/Month FE	✓	✓	✓	✓	✓	✓
MSA FE	✓	✓	✓	✓	✓	✓
Climate controls		✓	✓		✓	✓
Other controls			✓			✓
Obs.	2,882	2,852	2,852	2,882	2,852	2,852
Adj. R^2	0.788	0.792	0.793	0.655	0.656	0.658
	30 Largest MSAs					
	NO_2 Monthly Mean			**NO_2 Monthly Maximum**		
Entry	-0.012 (0.036)	-0.014 (0.035)	-0.021 (0.036)	-0.041 (0.027)	-0.039 (0.026)	-0.039 (0.024)
Controls included:						
Year/Month FE	✓	✓	✓	✓	✓	✓
MSA FE	✓	✓	✓	✓	✓	✓
Climate controls		✓	✓		✓	✓
Other controls			✓			✓
Obs.	2,732	2,732	2,732	2,732	2,732	2,732
Adj. R^2	0.774	0.776	0.777	0.652	0.655	0.655

Table A3. TWFE Regression of Ride-hailing Entry on Air Quality using EPA $PM_{2.5}$ Data on High Density/Large MSAs

	4th Density Quartile MSAs					
	$PM_{2.5}$ Monthly Mean			**$PM_{2.5}$ Monthly Maximum**		
Entry	-0.097 **	-0.093 *	-0.063	-0.042	-0.041	-0.010
	(0.047)	(0.048)	(0.043)	(0.055)	(0.057)	(0.053)
Controls included:						
Year/Month FE	✓	✓	✓	✓	✓	✓
MSA FE	✓	✓	✓	✓	✓	✓
Climate controls		✓	✓		✓	✓
Other controls			✓			✓
Obs.	2,981	2,921	2,921	2,981	2,921	2,921
Adj. R^2	0.419	0.413	0.425	0.428	0.419	0.420

	30 Largest MSAs					
	$PM_{2.5}$ Monthly Mean			**$PM_{2.5}$ Monthly Maximum**		
Entry	-0.105 *	-0.106 **	-0.088 *	-0.083	-0.082	-0.053
	(0.051)	(0.051)	(0.049)	(0.074)	(0.074)	(0.068)
Controls included:						
Year/Month FE	✓	✓	✓	✓	✓	✓
MSA FE	✓	✓	✓	✓	✓	✓
Climate controls		✓	✓		✓	✓
Other controls			✓			✓
Obs.	2,526	2,526	2,526	2,526	2,526	2,526
Adj. R^2	0.400	0.401	0.413	0.400	0.403	0.409

Appendix E. Population Density and Public Transit Use

In the main manuscript, we use population density as a sufficient statistic for unobserved trends and estimate the TWFE/event-study regressions on each density cohort, assuming that the parallel trends in outcomes to hold within the same density cohort. The primary reasons for our choice of population density are that we retain sufficient variations across cohorts and that we have pre-treatment data that can be consistently applied across cities. For example, we would have only the fourth quartile against all other cities if we use the share of train use as of 2010, for then the 25th and 50th percentiles are zeros. An implicit assumption here is that population density is a good predictor of the change in transit usage or public transit access/richness. However, it is known that some U.S. cities with high density are not necessarily "transit-friendly". So two important questions arise: (1) Does the high population-density cohort include mostly transit-friendly cities? and (2) Is our result robust to alternative cohort definitions that are more closely related to public transit use/availability?

On the first question, Table A4 below report the population size, the share of public transit commuters, and the share of rail transit commuters for top 40 cities in the highest population density cohort. In our sample, cities with public transit commuter shares exceeding 2% are relatively transit-friendly while those with nonzero rail transit shares are rail-accessible cities. We see that the highest density quartile includes all cities that are typically considered "transit friendly" (e.g., New York, San Francisco, Boston, Washington, D.C., Philadelphia, Chicago, Seattle, New Jersey/Newark, Portland, Baltimore, and Minneapolis–St. Paul) so our results are likely to be driven primarily by these cities.

On the second question, we ran three sets of TWFE regressions per eq. (1) in the main manuscript, using the quartiles of (i) the MSA-level public transit density (using data on bus transit network as of 2021 and rail transit network as of 2012), (ii) the MSA-level average share of rail transit commuter as of 2010, and (iii) the MSA-level average share of but transit commuter as of 2010. Only (i) and (iii) give us distinct quartiles while (ii) give us the fourth quartile versus all other cities. The results are reported in Tables A5, A6 and A7, respectively. We see that our results are largely robust to alternative metrics --- the estimates on the fourth quartile are negative, statistically significant, and larger for urban areas. We do see, however, important differences that seem quite intuitive. The estimates using (i) public transit density are quite similar to our original estimates presumably because the two metrics are highly correlated. The estimates using (ii) the share of rail transit commuters are quite similar when 90-th percentile is used as cutoff, but get smaller in magnitude when 75-th percentile is used as cutoff. In contrast, the estimates get smaller when we use (iii) the share of bus commuters as a metric. In addition, the estimates using (ii) or (iii) are larger in urban/suburban areas than in non-urban areas. These results suggest that the pollution-reducing effect of ride-hailing is driven mainly by cities with rail transit in urban and suburban areas. These results are also consistent with Krishnamurthy and Ngo (2024) who find that ride-hailing tends to reduce congestion and improve traffic speed in areas where the pre-treatment levels of public transit use are high, using cross-county variation in California. In Table A8, we also report the result of the TWFE regression using the same population density cohort as in the main manuscript, but including the transit density quartile-month fixed effects. The transit density quartiles here are the same quartiles as (i)-(iii) above. Inclusion of such fixed effects should control for the seasonality of public transit usage, and hence, serves as an additional robustness check. The results are again robust to the inclusion of these fixed effects.

Table A4. Top Population Density MSAs

CBSA_name	Pop. Density (2010)	Population (2010)	Public Transit Share (2010)	Rail Transit Share (2010)
New York-Northern New Jersey-Long Island, NY-NJ-PA	2,826	18,897	27.00	3.98
Los Angeles-Long Beach-Santa Ana, CA	2,646	12,829	5.50	0.23
San Francisco-Oakland-Fremont, CA	1,755	4,335	14.00	0.83
Trenton-Ewing, NJ	1,632	367	7.75	4.18
Honolulu, HI	1,587	953	8.05	0.00
Bridgeport-Stamford-Norwalk, CT	1,467	917	7.77	5.23
New Haven-Milford, CT	1,427	862	3.76	0.89
Chicago-Joliet-Naperville, IL-IN-WI	1,315	9,461	10.78	3.46
Boston-Cambridge-Quincy, MA-NH	1,305	4,552	9.89	1.92
Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	1,296	5,965	7.59	2.22
Tampa-St. Petersburg-Clearwater, FL	1,107	2,783	1.30	0.01
Detroit-Warren-Livonia, MI	1,105	4,296	1.35	0.00
Miami-Fort Lauderdale-Pompano Beach, FL	1,096	5,565	2.77	0.17
Milwaukee-Waukesha-West Allis, WI	1,070	1,556	2.99	0.04
Baltimore-Towson, MD	1,042	2,710	4.97	0.79
Cleveland-Elyria-Mentor, OH	1,040	2,077	3.07	0.09
Providence-New Bedford-Fall River, RI-MA	1,009	1,601	2.49	1.03
Washington-Arlington-Alexandria, DC-VA-MD-WV	997	5,582	12.40	0.77
Hartford-West Hartford-East Hartford, CT	800	1,212	2.31	0.02
El Paso, TX	791	801	1.74	0.04
Cape Coral-Fort Myers, FL	789	619	0.91	0.00
Akron, OH	781	703	1.29	0.00
San Diego-Carlsbad-San Marcos, CA	736	3,095	2.92	0.29
Buffalo-Niagara Falls, NY	726	1,136	3.02	0.00
Dallas-Fort Worth-Arlington, TX	714	6,372	1.26	0.22
San Jose-Sunnyvale-Santa Clara, CA	686	1,837	2.56	0.77
Houston-Sugar Land-Baytown, TX	674	5,947	2.11	0.01
Flint, MI	668	426	0.64	0.00
Bremerton-Silverdale, WA	636	251	10.12	0.00
Virginia Beach-Norfolk-Newport News, VA-NC	636	1,672	1.30	0.00
Atlanta-Sandy Springs-Marietta, GA	632	5,269	2.76	0.12
Orlando-Kissimmee-Sanford, FL	614	2,134	1.39	0.01
Santa Cruz-Watsonville, CA	589	262	1.94	0.00
Racine, WI	588	195	0.82	0.16
Seattle-Tacoma-Bellevue, WA	586	3,440	7.71	0.40
Charlotte-Gastonia-Rock Hill, NC-SC	570	1,758	1.74	0.06
Allentown-Bethlehem-Easton, PA-NJ	565	821	1.33	0.06
Lancaster, PA	550	519	0.98	0.15
Barnstable Town, MA	548	216	1.79	0.10
Minneapolis-St. Paul-Bloomington, MN-WI	544	3,280	3.18	0.10

Table A5. TWFE Estimates of Effect of Ride-hailing Entry on Ambient Air Quality, using Public Transit Density Quartiles

	Mean NO_2 Concentration								
	Urban Area			Suburban Area			Non-urban Area		
Entry									
× 1st PT Quartile MSAs	-0.004	0.003	0.003	-0.002	0.005	0.006	0.015	0.024	0.025
	(0.013)	(0.014)	(0.014)	(0.013)	(0.013)	(0.014)	(0.015)	(0.016)	(0.017)
× 2nd PT Quartile MSAs	-0.014	-0.009	-0.008	-0.012	-0.008	-0.006	-0.004	0.002	0.005
	(0.014)	(0.014)	(0.014)	(0.012)	(0.012)	(0.013)	(0.011)	(0.011)	(0.011)
× 3rd PT Quartile MSAs	-0.022	-0.014	-0.011	-0.024	-0.016	-0.013	-0.024	-0.016	-0.011
	(0.014)	(0.014)	(0.013)	(0.014)	(0.014)	(0.013)	(0.013)	(0.013)	(0.012)
× 4th PT Quartile MSAs	-0.038 **	-0.033 *	-0.029 *	-0.037 **	-0.032 **	-0.029 *	-0.039 **	-0.035 **	-0.030 *
	(0.014)	(0.014)	(0.014)	(0.013)	(0.014)	(0.014)	(0.012)	(0.014)	(0.014)
Controls included:									
Year/Month FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
MSA FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Climate controls		✓	✓		✓	✓		✓	✓
Other controls			✓			✓			✓
Obs.	38,911	35,023	35,023	39,309	35,421	35,421	39,254	35,367	35,367
Adj. R^2	0.759	0.757	0.757	0.765	0.762	0.762	0.780	0.778	0.778

Note : Clustered robust standard errors in parenthesis.

	Max NO_2 Concentration								
	Urban Area			Suburban Area			Non-urban Area		
Entry									
× 1st PT Quartile MSAs	-0.007	-0.001	0.000	-0.005	0.001	0.002	0.009	0.017	0.019
	(0.016)	(0.016)	(0.016)	(0.015)	(0.015)	(0.015)	(0.012)	(0.011)	(0.013)
× 2nd PT Quartile MSAs	-0.026	-0.025	-0.023	-0.025	-0.024	-0.022	-0.015	-0.014	-0.011
	(0.020)	(0.020)	(0.021)	(0.019)	(0.019)	(0.019)	(0.014)	(0.013)	(0.015)
× 3rd PT Quartile MSAs	-0.039 **	-0.03	-0.027	-0.041 *	-0.032	-0.028	-0.042 **	-0.034 *	-0.029 *
	(0.017)	(0.017)	(0.015)	(0.018)	(0.018)	(0.016)	(0.016)	(0.016)	(0.015)
× 4th PT Quartile MSAs	-0.059 ***	-0.056 ***	-0.052 ***	-0.059 ***	-0.057 ***	-0.053 ***	-0.061 ***	-0.061 ***	-0.055 ***
	(0.014)	(0.013)	(0.012)	(0.014)	(0.013)	(0.013)	(0.013)	(0.012)	(0.013)
Controls included:									
Year/Month FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
MSA FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Climate controls		✓	✓		✓	✓		✓	✓
Other controls			✓			✓			✓
Obs.	38,911	35,023	35,023	39,309	35,421	35,421	39,254	35,367	35,367
Adj. R^2	0.662	0.660	0.661	0.679	0.677	0.677	0.712	0.709	0.709

Note : Clustered robust standard errors in parenthesis.

Table A6. TWFE Estimates of Effect of Ride-hailing Entry on Ambient Air Quality, using the Share of Rail Transit Commuters as of 2010

	Mean NO_2 Concentration								
	Urban Area			Suburban Area			Non-urban Area		
Entry									
× All Other MSAs	-0.016	-0.010	-0.008	-0.015	-0.010	-0.007	-0.011	-0.005	-0.001
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.010)	(0.011)	(0.011)
× 90th Percentile MSAs	-0.050 ***	-0.042 ***	-0.039 **	-0.049 ***	-0.041 ***	-0.039 **	-0.037 **	-0.027 **	-0.023 *
	(0.014)	(0.012)	(0.012)	(0.013)	(0.012)	(0.012)	(0.011)	(0.010)	(0.011)
Controls included:									
Year/Month FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
MSA FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Climate controls		✓	✓		✓	✓		✓	✓
Other controls			✓			✓			✓
Obs.	38,911	35,023	35,023	39,309	35,421	35,421	39,254	35,367	35,367
Adj. R^2	0.759	0.757	0.757	0.765	0.762	0.762	0.78	0.777	0.778

Note : Clustered robust standard errors in parenthesis.

	Max NO_2 Concentration								
	Urban Area			Suburban Area			Non-urban Area		
Entry									
× All Other MSAs	-0.028 *	-0.024	-0.021	-0.028 *	-0.025 *	-0.021	-0.025 *	-0.021 *	-0.017
	(0.014)	(0.013)	(0.013)	(0.014)	(0.013)	(0.013)	(0.011)	(0.010)	(0.011)
× 90th Percentile MSAs	-0.072 **	-0.064 **	-0.061 **	-0.071 **	-0.063 **	-0.060 **	-0.057 **	-0.047 **	-0.043 *
	(0.023)	(0.022)	(0.021)	(0.024)	(0.022)	(0.022)	(0.021)	(0.019)	(0.019)
Controls included:									
Year/Month FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
MSA FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Climate controls		✓	✓		✓	✓		✓	✓
Other controls			✓			✓			✓
Obs.	38,911	35,023	35,023	39,309	35,421	35,421	39,254	35,367	35,367
Adj. R^2	0.662	0.660	0.661	0.679	0.677	0.677	0.712	0.709	0.709

Note : Clustered robust standard errors in parenthesis.

Table A7. TWFE Estimates of Effect of Ride-hailing Entry on Ambient Air Quality, using the Share of Bus Transit Commuters as of 2010

	Mean NO_2 Concentration								
	Urban Area			**Suburban Area**			**Non-urban Area**		
Entry									
× 1st Bus Quartile MSAs	-0.007	-0.003	0.000	-0.009	-0.005	-0.002	-0.008	-0.004	0.001
	(0.013)	(0.013)	(0.012)	(0.012)	(0.012)	(0.011)	(0.012)	(0.012)	(0.011)
× 2nd Bus Quartile MSAs	-0.020	-0.015	-0.012	-0.019	-0.014	-0.011	-0.017	-0.013	-0.009
	(0.011)	(0.010)	(0.011)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.011)
× 3rd Bus Quartile MSAs	-0.022	-0.015	-0.013	-0.021	-0.014	-0.012	-0.014	-0.007	-0.003
	(0.014)	(0.014)	(0.014)	(0.012)	(0.013)	(0.013)	(0.011)	(0.011)	(0.012)
× 4th Bus Quartile MSAs	-0.028 **	-0.020 *	-0.018	-0.026 **	-0.018 *	-0.016	-0.016	-0.007	-0.004
	(0.010)	(0.009)	(0.010)	(0.010)	(0.009)	(0.010)	(0.010)	(0.010)	(0.011)
Controls included:									
Year/Month FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
MSA FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Climate controls		✓	✓		✓	✓		✓	✓
Other controls			✓			✓			✓
Obs.	38,911	35,023	35,023	39,309	35,421	35,421	39,254	35,367	35,367
Adj. R^2	0.759	0.757	0.757	0.764	0.762	0.762	0.780	0.777	0.778

Note : Clustered robust standard errors in parenthesis.

	Max NO_2 Concentration								
	Urban Area			**Suburban Area**			**Non-urban Area**		
Entry									
× 1st Bus Quartile MSAs	-0.019	-0.017	-0.013	-0.023	-0.022	-0.017	-0.02	-0.020*	-0.014
	(0.013)	(0.013)	(0.013)	(0.014)	(0.013)	(0.013)	(0.012)	(0.011)	(0.011)
× 2nd Bus Quartile MSAs	-0.025	-0.023	-0.019	-0.023	-0.021	-0.017	-0.022 *	-0.021 *	-0.016
	(0.015)	(0.015)	(0.015)	(0.015)	(0.014)	(0.015)	(0.011)	(0.011)	(0.012)
× 3rd Bus Quartile MSAs	-0.035 *	-0.029 *	-0.026	-0.036 *	-0.029 *	-0.026	-0.031 **	-0.024 *	-0.02
	(0.016)	(0.016)	(0.015)	(0.016)	(0.015)	(0.015)	(0.013)	(0.012)	(0.013)
× 4th Bus Quartile MSAs	-0.048 **	-0.043 **	-0.04 **	-0.047 **	-0.042 **	-0.039 **	-0.037 **	-0.03 *	-0.027 *
	(0.016)	(0.015)	(0.014)	(0.016)	(0.015)	(0.014)	(0.015)	(0.013)	(0.013)
Controls included:									
Year/Month FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
MSA FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Climate controls		✓	✓		✓	✓		✓	✓
Other controls			✓			✓			✓
Obs.	38911	35023	35023	39309	35421	35421	39254	35367	35367
Adj. R^2	0.662	0.660	0.660	0.679	0.677	0.677	0.712	0.709	0.709

Note : Clustered robust standard errors in parenthesis.

Table A8. TWFE Estimates of Effect of Ride-hailing Entry on Ambient Air Quality, with Public Transit Density-Month Fixed Effects

	Mean NO_2 Concentration								
	Urban Area			Suburban Area			Non-urban Area		
Entry									
× 1st Quartile MSAs	0.004	0.009	0.009	0.005	0.01	0.01	0.029*	0.034	0.034
	(0.014)	(0.017)	(0.017)	(0.013)	(0.016)	(0.016)	(0.015)	(0.019)	(0.019)
× 2nd Quartile MSAs	-0.011	-0.010	-0.009	-0.010	-0.009	-0.009	-0.006	-0.005	-0.005
	(0.012)	(0.012)	(0.011)	(0.011)	(0.011)	(0.011)	(0.010)	(0.010)	(0.010)
× 3rd Quartile MSAs	-0.003	-0.004	-0.003	-0.004	-0.005	-0.004	-0.003	-0.004	-0.004
	(0.014)	(0.014)	(0.014)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.012)
× 4th Quartile MSAs	-0.031 *	-0.032 **	-0.033 **	-0.030 *	-0.031 **	-0.032 **	-0.029 *	-0.030 **	-0.031 **
	(0.014)	(0.014)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.012)
Density-Month FE:									
Public Transit Density	✓			✓			✓		
Rail Commuter Share		✓			✓			✓	
Bus Commuter Share			✓			✓			✓
Obs.	35,023	35,023	35,023	35,421	35,421	35,421	35,367	35,367	35,367
Adj. R^2	0.772	0.761	0.76	0.777	0.766	0.765	0.795	0.78	0.779

Note : Clustered robust standard errors in parenthesis.

	Max NO_2 Concentration								
	Urban Area			Suburban Area			Non-urban Area		
Entry									
× 1st Quartile MSAs	-0.011	-0.005	-0.005	-0.010	-0.004	-0.004	0.013	0.019	0.018
	(0.018)	(0.018)	(0.018)	(0.016)	(0.017)	(0.017)	(0.014)	(0.016)	(0.015)
× 2nd Quartile MSAs	-0.023	-0.022	-0.022	-0.024	-0.023	-0.023	-0.020	-0.019	-0.020
	(0.019)	(0.019)	(0.019)	(0.020)	(0.019)	(0.019)	(0.016)	(0.015)	(0.015)
× 3rd Quartile MSAs	-0.019	-0.020	-0.019	-0.020	-0.021	-0.020	-0.020	-0.021	-0.021
	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.012)	(0.012)	(0.012)
× 4th Quartile MSAs	-0.046 **	-0.047 **	-0.048 **	-0.046 **	-0.046 **	-0.048 **	-0.046 **	-0.048 **	-0.049 **
	(0.017)	(0.017)	(0.017)	(0.016)	(0.016)	(0.016)	(0.014)	(0.014)	(0.015)
Density-Month FE:									
Public Transit Density	✓			✓			✓		
Rail Commuter Share		✓			✓			✓	
Bus Commuter Share			✓			✓			✓
Obs.	35,023	35,023	35,023	35,421	35,421	35,421	35,367	35,367	35,367
Adj. R^2	0.673	0.665	0.663	0.691	0.681	0.679	0.726	0.712	0.711

Note : Clustered robust standard errors in parenthesis.

Appendix F. Orthogonality of De Facto Entry Timing

Our empirical analysis relies on two identifying assumptions. The first is the conditional parallel trend, which roughly requires that *de facto* entry timing of ride-hailing is uncorrelated with trends in unobservables once we control for city-level pre-treatment covariates. The second is the orthogonality of our IVs to the *de facto* entry timing conditional on such covariates. As an empirical support for these assumptions, we run the following OLS and Tobit regressions:

$$EntryDate_c = \alpha + X_{c,2010}'\beta + Z_c'\gamma + \epsilon_c$$

where Y_c is either the indicator of entry or the de facto entry timing, X_c are MSA-level socioeconomic variables as of year 2010, and Z_c are Uber's official entry timing interacted with three geography-based instruments. For the Tobit regression, we assume the entry dates are truncated at December 2020 for no-entry MSAs. The results are reported in **Table A9**. The table confirm Hall *et al.*'s argument that ride-hailing companies entered cities mostly in the rank order of city's population or economic size, and nothing else. The results also suggest that Uber's official entry timing (or IVs) do not predict *de facto* entry timing.

Table A9. Association between entry timing and Pre-treatment Covariates

	1(Ever Enter)		De Facto Entry Date							
	OLS		OLS				Tobit			
Pop$_{2010}$	0.026	0.034	-0.617	-0.685 *	-0.563	-0.606	-2.827	-3.661 *	-3.505 *	-4.507 **
	(0.021)	(0.021)	(0.380)	(0.376)	(0.381)	(0.378)	(1.974)	(1.984)	(2.105)	(2.124)
Pop. density$_{2010}$	0.117 ***	0.129 ***	0.436	0.106	0.123	-0.074	-10.502 ***	-11.893 ***	-11.592 ***	-12.902 ***
	(0.033)	(0.033)	(0.540)	(0.612)	(0.550)	(0.589)	(3.185)	(3.271)	(3.420)	(3.421)
Median age$_{2010}$	0.003	0.003	-0.030	-0.056	-0.013	-0.044	-0.263	-0.294	0.004	-0.032
	(0.006)	(0.006)	(0.117)	(0.107)	(0.124)	(0.118)	(0.580)	(0.584)	(0.612)	(0.625)
Median income$_{2010}$	0.254	0.347	-2.689	-2.334	-2.051	-1.598	-24.407	-32.445 *	-17.807	-27.123
	(0.204)	(0.212)	(3.513)	(3.689)	(4.008)	(4.138)	(18.990)	(19.633)	(20.047)	(20.539)
College$_{2010}$	0.077	0.107	-3.584	4.333	-2.630	4.373	-10.341	-7.377	-14.436	-13.479
	(0.529)	(0.557)	(7.536)	(8.181)	(7.886)	(8.782)	(48.685)	(51.654)	(50.258)	(53.655)
Non-white$_{2010}$	0.222	0.214	-4.214	-4.746	-3.795	-5.117	-23.954	-23.570	-25.523	-26.208
	(0.229)	(0.231)	(2.978)	(3.023)	(3.332)	(3.455)	(21.665)	(21.790)	(23.130)	(23.700)
Manuf. emp. share$_{2010}$	0.764 **	0.728 **	-6.167	-7.934	-9.026	-10.044 *	-75.741 **	-73.591 **	-72.143 **	-67.480 **
	(0.323)	(0.336)	(5.598)	(5.279)	(5.602)	(5.511)	(30.178)	(31.089)	(32.779)	(33.290)
Poverty$_{2010}$	0.227	0.475	27.121	27.205	33.585	35.344	2.996	-18.643	52.110	26.575
	(1.258)	(1.267)	(25.371)	(26.155)	(29.313)	(29.698)	(119.220)	(120.079)	(128.989)	(129.353)
Car commuter$_{2010}$		-0.315		30.392 *		28.179		49.750		58.528
		(0.861)		(16.340)		(17.563)		(81.962)		(96.120)
Pub. transit commuter$_{2010}$		-2.361 **		31.623		27.323		234.579 **		275.069 **
		(1.188)		(25.477)		(25.803)		(111.734)		(130.369)
Official Entry Date ×Planned Route IV					0.000	-0.000			0.010	0.009
					(0.002)	(0.002)			(0.007)	(0.007)
Official Entry Date ×Historical Route IV					0.002	0.001			0.000	-0.002
					(0.001)	(0.001)			(0.007)	(0.008)
Official Entry Date ×Inconsequential Unit IV					-0.001	-0.001			0.001	-0.000
					(0.001)	(0.001)			(0.006)	(0.006)
Obs.	345	345	284	284	269	269	345	345	329	329

Appendix G. Geographic Distribution of Geography-based Instruments

We use three types of geography-based instruments commonly used in the empirical economic geography literature [see Redding and Turner (2015) for an excellent review]. The first is based on the highway construction plan as of 1947 developed under the mandate to serve military services. This 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 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. 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. 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. 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 following maps illustrate the entry timing and location using our *de facto* entry measure as well as how they compare with our geography-based instruments. Figure A13 shows the geographic distribution of our *de facto* entry measure and the urban areas used to calculate urban-area NO_2 concentrations in our main analysis. Figure A14 is a map showing which MSAs were planned to connect on the 1947 highway plan. As Duranton and Turner (2011; 2012) note, many interstate highways were built subsequently based on this highway plan. Figure A15 displays historical railroad maps between 1840 and 1970. 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. Figure A16 depicts the Euclidean spanning tree network using the cities that existed as of 1860 and had a population of 10,000 or more. 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.

Table A10 also reports OLS regressions of several measures of public transit density and commuter shares on the geography-based IVs. The results show that these IVs are associated with pre-treatment measures of public transit availability and use, although the estimates for public transit shares are imprecise and not always statistically significant, partly because the shares are close to zero and display limited variation in many cities.

Figure A13. Geographic Distribution of Ride-hailing Entry and Urban Area Boundaries

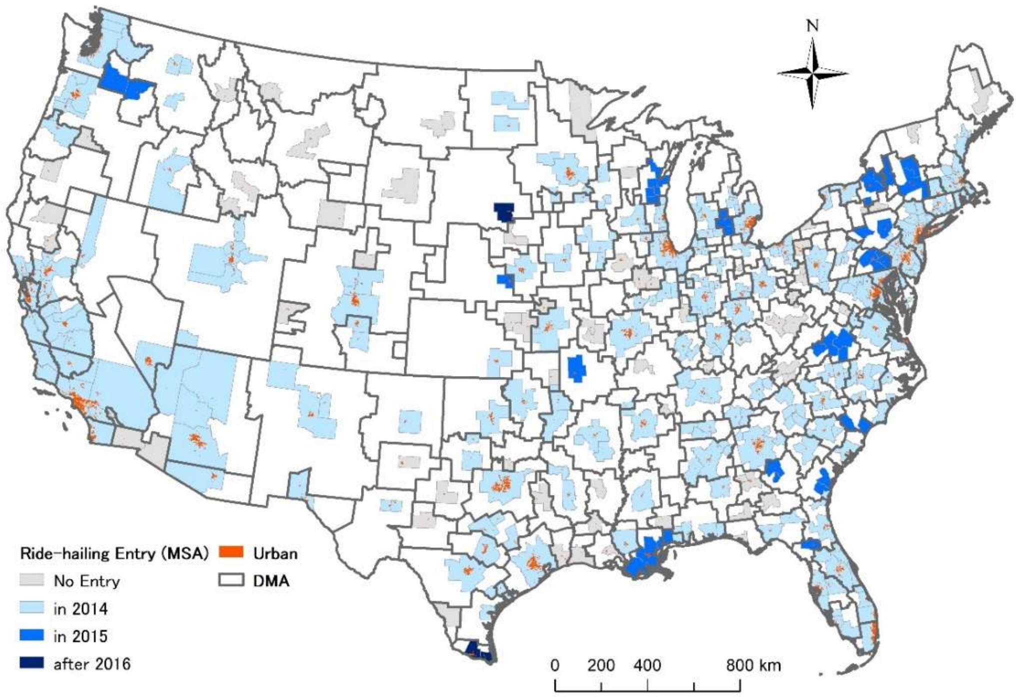


Figure A14. 1947 highway plan

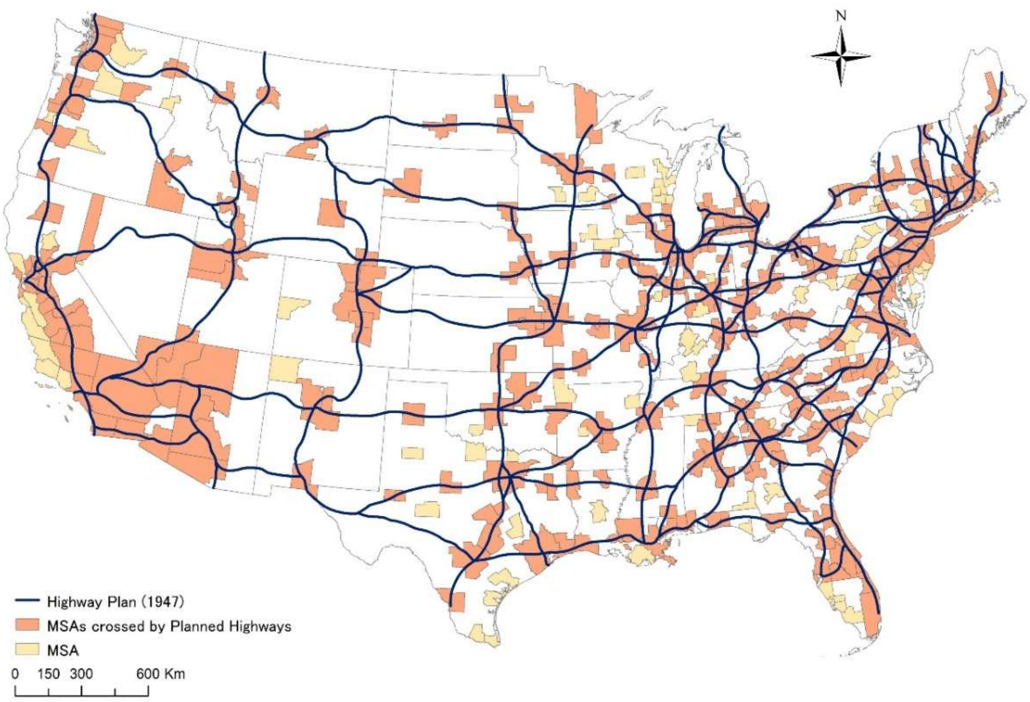


Figure A15. Railroad as of 1870

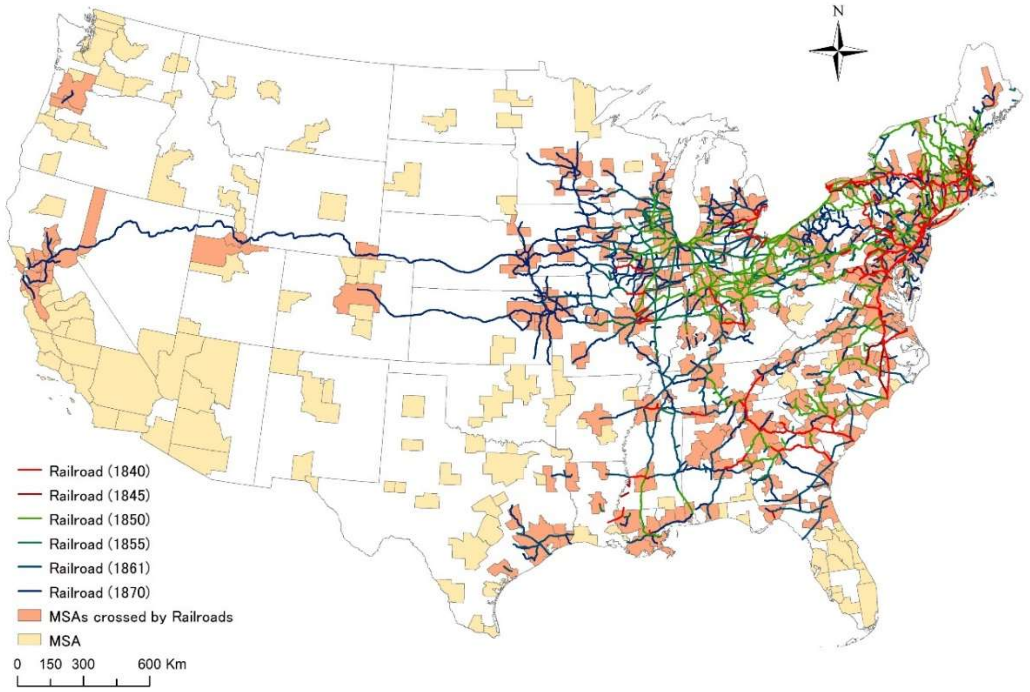


Figure A16. Euclidean network connecting large cities as of 1860

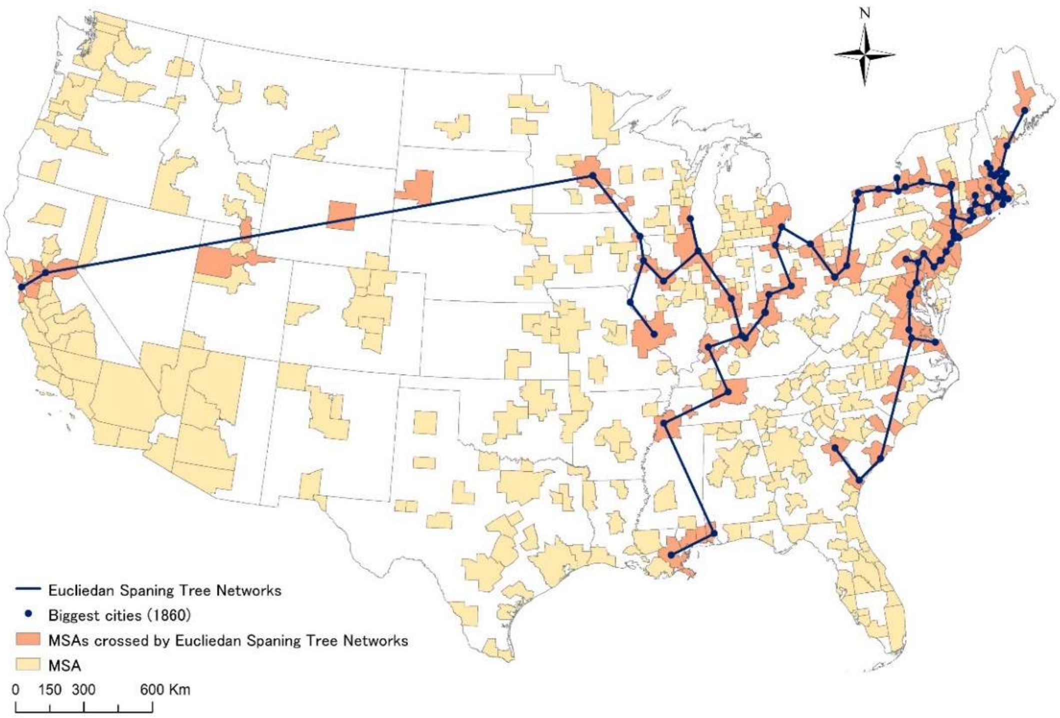


Table A10. Association between Geography-based IVs and Pre-treatment Public Transit Measures

	Transit Density		Pop. $Density_{2010}$		Car $commuter_{2010}$		Pub. transit $commuter_{2010}$	
Planned Route IV	0.006	0.003	0.158	-0.071	0.008	0.011 ***	0.001	-0.002
	(0.005)	(0.005)	(0.111)	(0.084)	(0.005)	(0.004)	(0.003)	(0.002)
Historical Route IV	0.021 ***	0.019 ***	0.449 ***	0.5 ***	0.022 ***	0.018 ***	-0.002	0.002
	(0.005)	(0.005)	(0.101)	(0.080)	(0.005)	(0.004)	(0.003)	(0.002)
Inconsequential Unit IV	0.021 ***	0.004	0.523 ***	0.032	-0.019 ***	-0.001	0.015 ***	0.003
	(0.006)	(0.006)	(0.117)	(0.093)	(0.006)	(0.005)	(0.003)	(0.003)
Pop_{2010}		0.008 ***		0.427 ***		-0.001		0.006 ***
		(0.003)		(0.041)		(0.002)		(0.001)
Median age_{2010}		0.001 **		0.045 ***		0.001 **		0.000
		(0.001)		(0.010)		(0.000)		(0.000)
Median $income_{2010}$		0.103 ***		1.431 ***		-0.051 ***		0.051 ***
		(0.024)		(0.399)		(0.020)		(0.012)
$College_{2010}$		0.068		3.116 ***		-0.323 ***		0.074 ***
		(0.057)		(0.944)		(0.047)		(0.027)
$Non\text{-}white_{2010}$		-0.027		0.535		0.043 **		-0.004
		(0.022)		(0.357)		(0.018)		(0.010)
Manuf. emp. $share_{2010}$		0.084 **		3.252 ***		0.072 **		-0.013
		(0.036)		(0.601)		(0.030)		(0.017)
$Poverty_{2010}$		0.274 **		5.876 ***		-0.112		0.146 **
		(0.136)		(2.237)		(0.110)		(0.065)
Obs.	345	345	345	345	345	345	345	345
F-stat on Excluded IV	16.78 ***	6.96 ***	21.52 ***	15.17 ***	9.27 ***	10.82 ***	9.27 ***	1.19

Appendix H. Binary vs. Continuous Treatment

In the main text, we estimated the effect of ride-hailing entry on air quality using a binary treatment indicator defined in Section 3. Although there are a number of advantages to doing so (as discussed in the main text), there is also one important disadvantage --- we fail to capture a dose-response relationship between the *intensity* of ride-hailing activity and air pollution. Ideally, we would estimate an elasticity of air pollution with respect to some measure of ride-hailing intensity such as the ride-hailing trip volume or the number of ride-hailing providers (adjusted for the size of city). However, we face a number of empirical limitations, which prevented us from doing so.

First, to our knowledge, there is no publicly or commercially available data on ride-hailing trip volume, market shares, or the number of ride-hailing drivers that can be used consistently over time across cities. Second, even if such data were available, they are still equilibrium outcomes --- how the rideshare intensity changes over time across cities is highly endogenous. Third, to address this endogeneity and to estimate the elasticity, we would need an exogenous instrument that would predict the level of rideshare intensity, not just the timing and location of entry. Fourth, as we argue in the paper, the elasticity of air pollution with respect to ride-hailing intensity depends critically on the complementarity between rideshare and public transit. This means that we would need exogenous sources of variation for both rideshare and public transit over time and across cities. These are precisely the reasons why previous studies such as Hall *et al.* (2018), Tarduno (2021), and Krishnamurthy and Ngo (2024) have relied mostly on the binary treatment like ours.

Despite these limitations, we also attempt to estimate the elasticity of the elasticity of air pollution with respect to the (combined) Google Trends in a manner analogous to Hall *et al.* (2018). Unlike Hall *et al.*, however, we use geography-based instruments to address the endogeneity of Google Trends index. Specifically, we estimate the following regression,

$$ln(y_{ctm}^{j,s}) = \alpha_c + \lambda_{tm} + \beta \ln(RS_{ctm}) + X'_{ctm}\gamma + \theta_c(t) + \epsilon_{ctm},$$

where RS_{ctm} is the Google Trends index defined as

$$Trend_{ctm} = \max\{Trend_{ctm}^{Uber}, Trend_{ctm}^{Lyft}\}.$$

We estimate this regression separately for each population density quartile $s \in S$ and each within-city area $j \in \{\text{urban, suburban, non-urban}\}$. By construction, the Google Trends index is normalized for the size of city so it is a normalized measure of ride-hailing intensity that can be consistently compared over time and across cities (Note that any measure of ride-hailing intensity such as rideshare trip volume and the number of ride-hailing providers need to be normalized against the size of the city such as population and the number of commuters). However, the index is not necessarily a good proxy to ride-hailing trip volume or the number of ride-hailing providers. It is also highly endogenous, as explained in the main text. To instrument $\ln(RS_{ctm})$, we construct a set of IVs by interacting our geography-based instruments with Uber's official entry timing. The estimates are reported in Table A11, along with weak IV and overidentification test statistics. We see that the results are consistent with those in the main text --- the estimated elasticity statistically significant only in the highest density quartile and ranges from -0.050 to -0.063, implying that one percent increase in market penetration leads to a decrease in NO_2 concentration levels by 0.05 to 0.063 percent. We, however, take these results as only suggestive, for all the reasons explained above.

Table A11. DD-IV Estimates of Elasticity of Air Quality with Ride-hailing Intensity

	Population Density				
	Lowest MSAs	**2nd MSAs**	**3rd MSAs**	**Highest MSAs**	
Monthly Mean NO_2					
Urban area	-0.014	-0.034	-0.001	-0.057	**
	(0.023)	(0.033)	(0.018)	(0.026)	
Suburban area	-0.014	-0.038	-0.005	-0.05	**
	(0.022)	(0.038)	(0.017)	(0.025)	
Non-urban area	-0.01	-0.036	-0.017	-0.058	**
	(0.022)	(0.043)	(0.017)	(0.024)	
Monthly Maximum NO_2					
Urban area	0.004	0.002	0.011	-0.054	
	(0.028)	(0.026)	(0.023)	(0.032)	
Suburban area	0.002	-0.015	-0.002	-0.056	*
	(0.026)	(0.027)	(0.021)	(0.032)	
Non-urban area	-0.001	-0.015	-0.027	-0.063	**
	(0.021)	(0.031)	(0.019)	(0.029)	
Obs.	8,460	8,601	8,762	8,936	
Weak IV stat.	111.467	31.283	98.454	122.549	
Hansen's J stat.	0.094	0.431	3.641	0.149	
(p-values)	0.954	0.806	0.162	0.928	

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

Appendix I. Full IV Regression Results on the Mediating Effect of Commuting Mode Choice: Instrument Relevance, Alternative Specifications, and Alternative Instruments

Instrument Relevance by Density Quartile: In Table 3 in the main text, we report the first-stage results using the full sample. Table A12 below reports the results by density strata. The results indicate that there is indeed sufficient within-strata variation so the F-stats of excluded instruments are sufficiently high particularly on the top quartile (highest density). We seem to lose within-strata variation for less dense cities. However, the direction, magnitude, and statistical significance of IV coefficients are similar between the top quartile and the full sample. Thus, much of our results in the mediation analysis comes from variation in the top density quartile, and this is also consistent with our main regression results.

Table A12. Instrument Relevance by Density Quartile

	Commuting by Car				Commuting by Public Transit			
	Lowest	2nd	3rd	Highest	Lowest	2nd	3rd	Highest
L.O.O Regional Exposure	--	-3127.613	-166.133	-807.261 **	--	-25.030	-61.347	107.359 *
×Planned Route IV	--	(2157.910)	(570.063)	(348.121)	--	(24.331)	(43.023)	(69.120)
L.O.O Regional Exposure	12959.23 ***	2763.191	-797.727	388.981	125.055 ***	4.838	-92.485	-77.617
×Historical Route IV	(307.939)	(2194.455)	(630.132)	(326.527)	(10.164)	(29.163)	(60.726)	(67.209)
L.O.O Regional Exposure	1045.608 ***	932.228 ***	1887.473 ***	544.311 ***	31.791 ***	35.52 *	186.06	127.362 ***
×Inconsequential Unit IV	(175.657)	(256.892)	(548.816)	(35.171)	(6.934)	(17.619)	(55.803)	(33.016)
F-statistic	8611.22 ***	10.87 ***	9.62 ***	122.74 ***	1722.17 ***	9.97 ***	9.88 ***	173.86 ***
Obs.	1,137	2,225	3,579	15,962	1,137	2,225	3,579	15,962

Estimation Results of Alternative Specifications: In addition, in Figure 7 of the main text, we only reported the coefficient estimates of the mediating effect from our preferred specification. Table A13 below reports the results of all regressions with varying levels of controls, along with the full first-stage statistics. For each outcome (monthly mean NO_2 or monthly maximum NO_2), we run three regressions with varying levels of controls, with and without IVs. The first column controls for MSA, year, and month fixed effects. This will control for all time-invariant covariates at the MSA level as well as time-varying confounds that are common to all units. The second column control for PUMA-level averages of household-level demographics such as income, education, and racial composition, which vary over time and across MSAs/PUMAs. This is our preferred specification and is used in Figure 7 of the main text. Ideally, one might wish to include PUMA-year fixed effects to account for arbitrary PUMA-specific shocks. In our setting, however, inclusion of PUMA-year fixed effects would absorb not only the PUMA-year demographic controls but also the treatment variables themselves. Inclusion of PUMA-level demographic controls, therefore, is the most feasible compromise for controlling for PUMA-specific time-varying confounds while preserving the identifying variation necessary for estimation. In this sense, the preferred specification offers the closest feasible approximation to a PUMA-year fixed-effects design. The third specification further includes MSA-by-year fixed effects. These fixed effects control for ride-hailing entry status as well as other MSA-specific time-varying confounds such as changes in industry composition, transportation networks, and local economic conditions. This specification is quite demanding, however: because the preferred specification already adjusts for the main observed sources of PUMA-level time-varying heterogeneity, adding MSA-year fixed effects may absorb substantial variation relevant for identification. We, therefore, view this as a robustness check rather than the preferred specification.

For all regressions, we report robust standard errors clustered at the MSA and PUMA level in parenthesis. As shown in the table, the estimates are quite robust to varying sets of controls. The table also reports the Kleibergen-Paap rk Wald F-statistic (under clustered standard errors and the Cragg-Donald F-statistic (under homoskedasticity) for the weak IV diagnostic as well as Hansen's J statistic for the test of overidentification.

Table A13. OLS and IV Estimates of Mediating Effect of Commuting Mode Choice

	Monthly Mean NO_2						Monthly Maximum NO_2					
	OLS			IV			OLS			IV		
Car Commuter	0.005 ***	0.005 ***	0.001 ***	0.015 ***	0.013 ***	0.013 ***	0.004 ***	0.004 ***	0.001 *	0.014 ***	0.012 ***	0.012 ***
	(0.001)	(0.001)	(0.000)	(0.003)	(0.003)	(0.003)	(0.001)	(0.001)	(0.000)	(0.003)	(0.003)	(0.003)
Public Transit Commuter	-0.002	-0.002 *	0.000	-0.007 **	-0.007 **	-0.007 **	-0.003	-0.003	-0.002	-0.005 *	-0.006 **	-0.006 **
	(0.002)	(0.001)	(0.000)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.001)	(0.003)	(0.003)	(0.003)
Controls included:												
MSA	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Year/Month	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Household Demographics		✓	✓		✓	✓		✓	✓		✓	✓
MSA-Year			✓			✓			✓			✓
Obs.	23,112	23,112	23,112	22,899	22,899	22,899	23,112	23,112	23,112	22,899	22,899	22,899
Weak IV F-stat.												
Kleibergen-Paap				14.961 ***	9.678 ***	9.778 ***				14.961 ***	9.678 ***	9.778 ***
Cragg-Donald				906.09 ***	797.01 ***	835.69 ***				906.09 ***	797.01 ***	835.69 ***
Hansen's J stat.				0.079	0.467	0.483				0.014	0.24	0.247
(p-values)				0.779	0.495	0.487				0.904	0.624	0.619

Alternative Instruments: In the main text, we use the interaction of the leave-one-out regional average exposure duration $\bar{\omega}^{o}_{r(c)}$ and the geography-based instruments G_c as our IVs:

$$Z_{cj} = \overline{\omega}^{o}_{r(c)} \times \frac{G_c}{A_{cj}},$$

in the spirit of Bartik shift-share instruments. While the geography-based instruments (the *share* component) are well established in the literature, the leave-one-out regional exposure duration (the *shift* component) is new to the literature. Although we take substantial care in designing this instrument, it is worth investigating alternative instruments for the shift component. Below, we explore three alternative instruments and compare the regression results across them.

The first obvious candidate for the shift component is to use the regional diffusion profile $\bar{\theta}^{o}_{r(c)t}$ in place of $\overline{\omega}^{o}_{r(c)}$:

$$\bar{\theta}^{o}_{r(c)t} = \frac{1}{n_{r(c)} - 1} \sum_{i \neq c, i \in C_{r(c)}} D^{\text{Uber}}_{it}$$

where D^{Uber}_{it} is an indicator for Uber's official presence in city i in year t. Again, we calculate this rollout profile at the Census region level, exclude the city's own contribution, and use Uber's official entry timing, for all the reasons discussed in the main text. Note that if we calculate this at the national level, it would essentially collapse into a year dummy in our 2-by-2 setup, making it virtually perfectly collinear with the city-year fixed effect.

The second option is to use the local exposure duration ω_c:

$$\omega_c = 2016 - E^{\text{Uber}}_{c},$$

where E^{Uber}_{c} denotes Uber's official entry year in city c. In fact, this was our initial attempt. Because Uber's administrative entry is a noisy measure of ride-hailing presence, this may also create counterfactual differential exposure duration in a manner similar to $\overline{\omega}^{o}_{r(c)}$. The drawback, however, is that Uber's rollout is influenced by city-specific conditions, and if so, ω_c may be correlated with unobservable determinants of commuting behavior or pollution. Our leave-one-out regional exposure instrument helps us mitigate such concerns, and hence, is our preferred instrument.

The third option is to remove the shift component altogether. That is, we simply use G_c/A_{cj} as our instrument. This helps us assess the relative contribution of the timing component.

Table A14 below reports the results of IV regressions using these alternative instruments, along with all relevant first-stage statistics. All regressions use the full set of controls (i.e., the same as the third columns of Table A13). For each outcome, the first column reports the result of our preferred instrument using $\bar{\omega}^{o}_{r(c)}$, the second column uses $\bar{\theta}^{o}_{r(c)t}$, the third uses ω_c, and the last column uses the geography-only instrument. In addition to the first-stage statistics, the table reports the F-statistics on the placebo regressions similar to Table 3-A.

Table A14. Estimation Results with Alternative Instruments

	Monthly Mean NO_2				**Monthly Maximum NO_2**			
	IV				IV			
	L.O.O Exposure	Regional Rollout	Local Exposure	Geography Only	L.O.O Exposure	Regional Rollout	Local Exposure	Geography Only
Car Commuter	0.013 ***	0.016 ***	0.013 ***	0.013 ***	0.012 ***	0.015 ***	0.013 ***	0.012 ***
	(0.003)	(0.004)	(0.002)	(0.003)	(0.003)	(0.004)	(0.002)	(0.003)
Public Transit Commuter	-0.007 **	-0.009 **	-0.008 ***	-0.007 **	-0.006 **	-0.007 **	-0.008 ***	-0.006 **
	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Controls included:								
MSA	✓	✓	✓	✓	✓	✓	✓	✓
Year/Month	✓	✓	✓	✓	✓	✓	✓	✓
Household Demographics	✓	✓	✓	✓	✓	✓	✓	✓
MSA-Year	✓	✓	✓	✓	✓	✓	✓	✓
Obs.	22,899	23,100	22,899	23,100	22,899	23,100	22,899	23,100
Weak IV F-stat.								
Kleibergen-Paap	9.778 ***	8.749 ***	9.579 ***	9.182 ***	9.778 ***	8.749 ***	9.579 ***	9.182 ***
Cragg-Donald	835.687 ***	267.443 ***	709.548 ***	839.698 ***	835.687 ***	267.443 ***	709.548 ***	839.698 ***
Hansen's J stat.	0.483	1.015	0.123	0.436	0.247	0.882	0.123	0.209
(p-values)	0.487	0.314	0.726	0.509	0.619	0.348	0.726	0.648
Placebo F-stat.	1.868	5.145 ***	1.930	1.782	1.271	5.743 ***	0.970	0.983

As shown in the table, our main results are robust to alternative instruments, and the main estimates and the first-stage statistics are quantitatively quite similar across alternative instruments. There is one important exception, however: when the regional diffusion profile $\bar{\theta}^{o}_{r(c)t}$ is used, the placebo regressions become statistically significant and first-stage strength weakens somewhat. We interpret this pattern as evidence that the regional diffusion profile may be somewhat mechanically correlated with contemporaneous city-level ride-hailing penetration than our preferred leave-one-out regional exposure-duration measure. For this reason, we prefer the latter as the shift component. Furthermore, we see that our regional-exposure IV marginally improves on the first-stage statistics, relative to either the local-exposure IVs or the geography-only IVs, while leaving the placebo F-statistics virtually intact. Taken together, these results suggest that the estimated impacts are robust to alternative instruments, and reinforce the argument that our IVs provide the cross-sectional and temporal variation necessary for identification while remaining plausibly exogenous to contemporaneous city-level shocks, conditional on included covariates.

Appendix J. Effect of Ride-hailing Entry on Commuting Mode Choice using Full Sample

In the main text, we conducted our causal mediation analysis by restricting our study sample to 2012 and 2016. This was done mainly for the clear identification of the effect of our mediator, commuting mode choice. There is nothing wrong to use the full sample, however, if the sole purpose was to estimate the effect of ride-hailing on commuting mode choice. Use of the full sample also allows us to address the concerns about pre-trends.

Table A15-A reports the TWFE estimates of the effect of ride-hailing entry on commuting mode choice using the full sample. As shown, the results are very similar to the ones using the restricted sample. In all regressions, we include the MSA and year fixed effects as well as household-level demographic controls. For each outcome, the second column includes weather controls whereas the third column also includes other time-varying controls such as gasoline price and nonattainment status.

Table A15-B repeats the same, but use the sub-categories of I_o as outcomes (I_{Walk}, I_{Bicyle}, I_{Home}, I_{Taxi}, $I_{MotorCycle}$, and I_{othe}). The table reveals something we did not expect a priori. In the highest density MSAs, ride-hailing entry is significantly associated with increases in commuting by walking (I_{Walk}), bicycling (I_{Bicyle}), and working at home (I_{Home}). We interpret the increase in walking and bicycling as the sign of the complementary use of ride-hailing alongside other commuting modes such as public transit while the rise in those working at home as the sign of the associated increase in ride-service providers.

Table A15. TWFE Estimates of Effect of Ride-hailing Entry on Commuting Mode Choice

A. Primary Commuting Mode

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

B. Other Commuting Mode: Subcategories

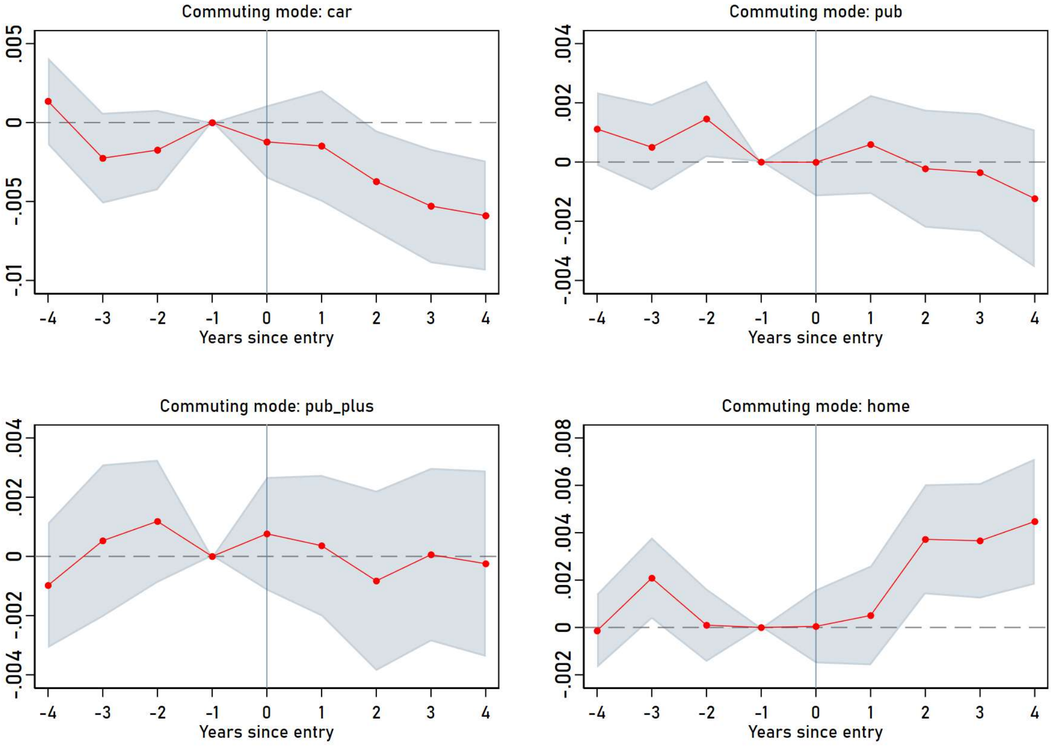
	Commuting Mode to Work					
	Walking	**Bicycling**	**Work Home**	**Taxi**	**Motor Cycle**	**Other**
Entry						
× 1st Quartile MSAs	-0.001	0.000	-0.001	0.000 **	0.000	-0.001
	(0.001)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)
× 2nd Quartile MSAs	-0.001 *	0.000	0.001	0.000	0.000	0.000
	(0.001)	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)
× 3rd Quartile MSAs	0.000	0.000 *	0.005 ***	0.000	0.000	0.000
	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)
× 4th Quartile MSAs	0.001 ***	0.001 ***	0.005 ***	0.000 ***	0.000 ***	0.000
	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)
Controls included:						
Year FE	✓	✓	✓	✓	✓	✓
MSA FE	✓	✓	✓	✓	✓	✓
Demog. Controls	✓	✓	✓	✓	✓	✓
Climate controls	✓	✓	✓	✓	✓	✓
Other controls	✓	✓	✓	✓	✓	✓
Mean in 2010	2.52%	0.51%	4.46%	0.10%	0.22%	0.74%
Obs.	17,872,388	17,872,388	17,872,388	17,872,388	17,872,388	17,872,388
Adj. R^2	0.007	0.004	0.01	0.002	0.001	0.002

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

Our TWFE/event-study regressions are intended to compare outcome trends of early treated cities against those of not-yet-treated cities within the same population density cohort. The identifying assumption here is that the cities within the same cohort would have similar outcome trends (i.e., commuting mode changes) in the absence of treatment. In other words, our empirical strategy is valid as long as outcome trends are similar within the same cohort, but not if they differ across cities. To examine if the parallel trends hold, we estimate event-study regressions using the dCDH estimator on the full sample from 2010 to 2018, employing the staggered DD design. If pre-trends systematically differ across cities, we should observe statistically significant pre-treatment placebo coefficients. Note, however, that unlike air pollution regressions, we had difficulty removing time-varying covariates a la Catenao *et al.* (2023). Thus, we report the unconditional event-study estimates in **Figure A17** for the full sample (Panel A) and the 4th density quartile (Panel B) using the *de facto* entry as the treatment variable. **Figure A18** reports the same using the refined entry defined in **Appendix C**. **Figure A17** and **A18** reveal that although standard errors are large, parallel trends seem to hold during the pre-treatment periods on the highest density cohort, so the pre-trends are unlikely to be a major concern.

Figure A17. Event-study Estimates of the Effect of *De Facto* Entry on Commuting Mode Shares

A. Full Sample



B. Highest Density Quartile

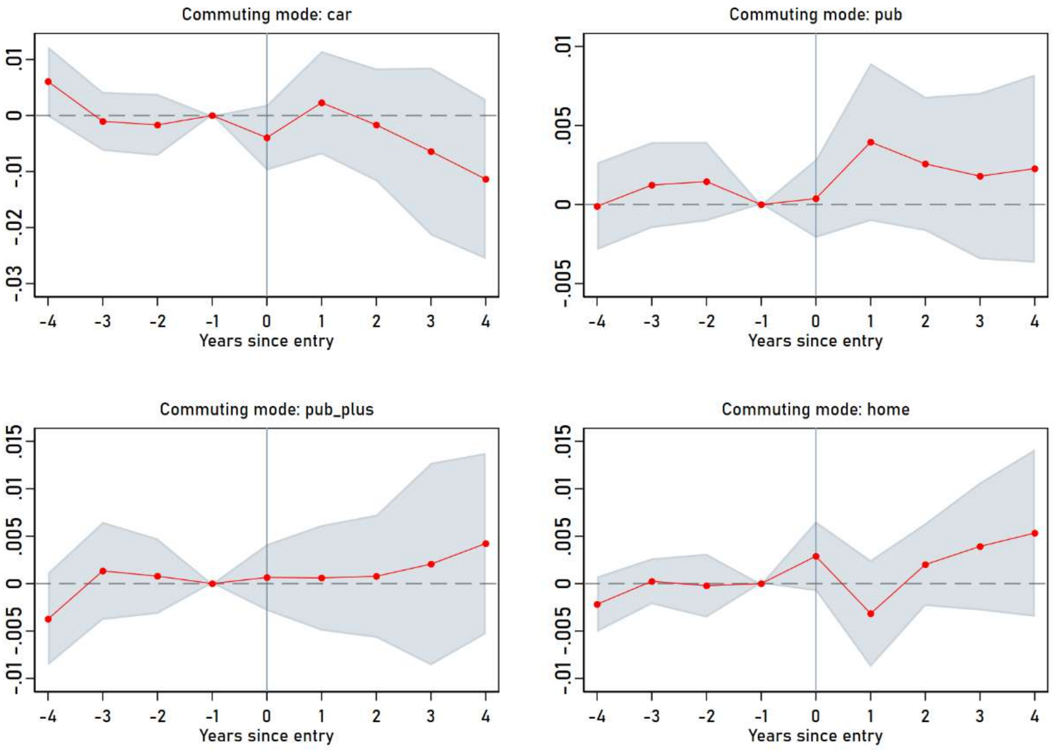
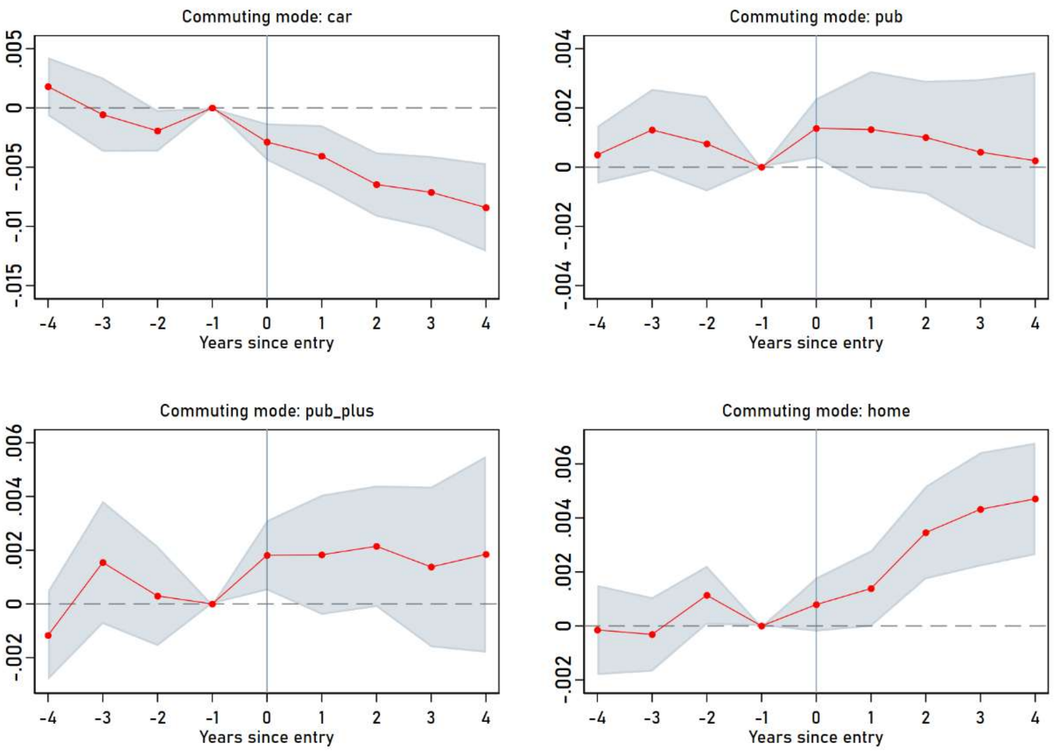
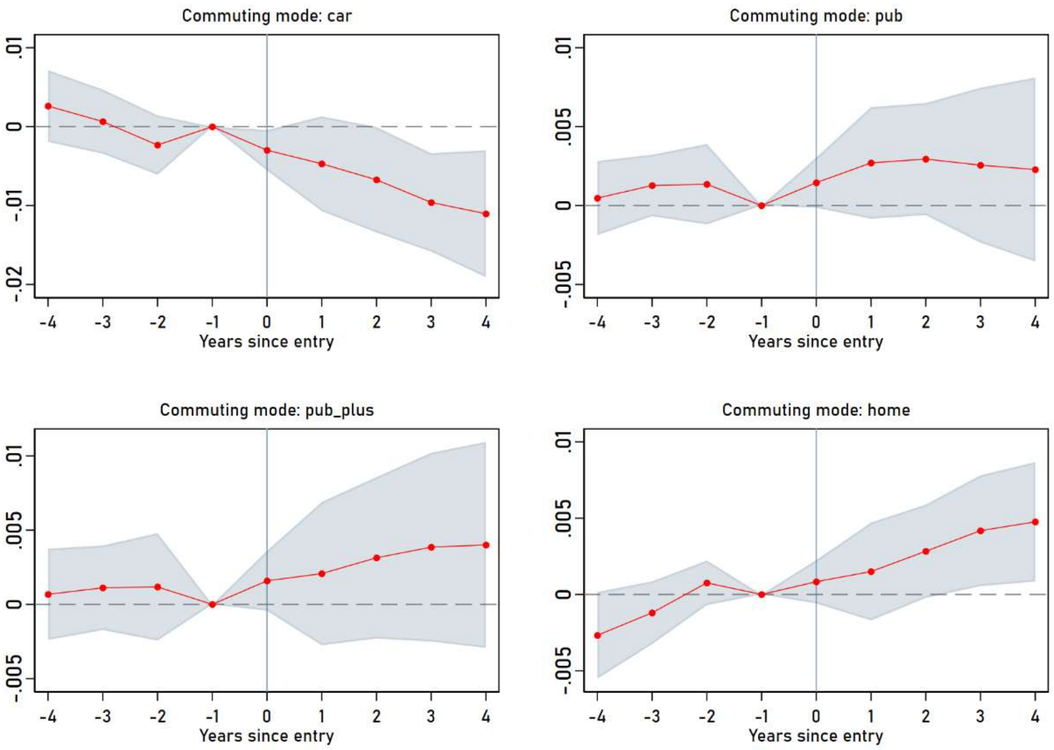


Figure A18. Event-study Estimates of the Effect of *Refined* Entry on Commuting Mode Shares

A. Full Sample



B. Highest Density Quartile



Appendix K. Implied Average Indirect Effects under Alternative Assumptions

Our estimates of the implied average indirect effect (AIE) reported in Table 4 of the main manuscript are somewhat conservative. In this appendix, we provide the estimates under several alternative assumptions.

First, the original estimates use the estimates of the impacts of ride-hailing entry on commuting modes from the 2-by-2 DD regressions on the restricted subsample. These estimates are smaller in magnitude than those of the TWFE regression under the staggered DD design on the full sample (see Panel A in Table A15): $\widehat{I(car)} = -0.0111$ vs. -0.0126 and $\widehat{I(pt)} = 0.0026$ vs. 0.0039 in the highest density cohort. Panel B of Table A16 below reports the estimates of the implied AIE using the estimates from the TWFE regressions with full controls reported in Panel A of Table A15.

Second, as shown in Appendix J, ride-hailing also increased the commuter share of walking and bicycling. These commuting modes are likely to be used in combination with ride-hailing, and thus, may have a similar pollution-reducing effect under certain realistic assumptions, as discussed in Appendix B. Panel C of Table A16 reports the estimates of the implied AIE incorporating the estimated changes in walking and bicycling reported in Panel B of Table A15.

Lastly, the original estimates are calculated as $\hat{\beta}_{car}\widehat{I(car)} + \hat{\beta}_{pt}\widehat{I(pt)}$. However, if we strictly follow the reasoning in Appendix B, we could estimate the implied effect as $\hat{\beta}_{car}\widehat{I(car)} + (\hat{\beta}_{pt} - \hat{\beta}_{car})\widehat{I(pt)}$. In our view, this estimate is too optimistic --- since the estimate $\hat{\beta}_{pt}$ is already negative, presumably it is already capturing the net effect of public transit relative to car commuting. Nonetheless, Panel D of Table A16 reports the estimates of the implied AIE using this latter formula.

Table A16 also compares the implied mediating effect of commuting mode change against the direct estimates of the effect of ride-hailing on monthly average NO_2 concentrations in a similar manner to Table 4. For ease of comparison, the original estimates are also shown in Panel A and the direct estimates are reported in the bottom row. To avoid a busy table, all these results are reported only for the highest density cohort.

As expected, the implied AIE estimates increase gradually from Panel A to Panel D. In particular, the estimates are remarkably close to the direct estimates when evaluated at the mean density of the top 30 PUMAs using the estimated commuting mode changes $\hat{I}$ on the highest density cohort.

Table A16. The Mediating Effect of Commuting Mode Change under Alternative Assumptions

				MSA-level Population Density Quartile				
				Lowest	2nd	3rd	Highest	Overall
A. 2-by-2 DD, Restricted Sample								
		ΔI_{car}		0.0014	-0.0040	-0.0057	-0.0111	-0.0084
		$\Delta I_{pub.\ transit}$		0.0004	-0.0004	0.0001	0.0026	0.0016
MSA-level		**PUMA-level**						
Highest	Mean		17.57	0.000	-0.001	-0.001	-0.003	-0.002
	Min		0.08	0.000	0.000	0.000	0.000	0.000
	Max		298.65	0.005	-0.014	-0.022	-0.048	-0.036
	Top 30		104.00	0.002	-0.005	-0.008	-0.017	-0.012
B. Staggered DD, Full Sample								
		ΔI_{car}		0.0031	0.0008	-0.0041	-0.0126	-0.0083
		$\Delta I_{pub.\ transit}$		-0.0009	-0.0011	-0.0007	0.0039	0.0021
MSA-level		**PUMA-level**						
Highest	Mean		17.57	0.001	0.000	-0.001	-0.003	-0.002
	Min		0.08	0.000	0.000	0.000	0.000	0.000
	Max		298.65	0.014	0.006	-0.014	-0.057	-0.036
	Top 30		104.00	0.005	0.002	-0.005	-0.020	-0.013
C. Staggered DD, Full Sample, Walking & Bicycling Included								
		ΔI_{car}		0.0031	0.0008	-0.0041	-0.0126	-0.0083
		$\Delta I_{pub\ tr.,\ walk,\ bicycle}$		-0.0015	-0.0018	-0.0003	0.0065	0.0036
MSA-level		**PUMA-level**						
Highest	Mean		17.57	0.001	0.000	-0.001	-0.004	-0.002
	Min		0.08	0.000	0.000	0.000	0.000	0.000
	Max		298.65	0.015	0.007	-0.015	-0.063	-0.040
	Top 30		104.00	0.005	0.003	-0.005	-0.022	-0.014
D. Staggered DD, Full Sample, Walking & Bicycling Included, Optimistic								
		ΔI_{car}		0.0031	0.0008	-0.0041	-0.0126	-0.0083
		$\Delta I_{pub\ tr.,\ walk,\ bicycle}$		-0.0015	-0.0018	-0.0003	0.0065	0.0036
MSA-level		**PUMA-level**						
Highest	Mean		17.57	0.001	0.001	-0.001	-0.005	-0.002
	Min		0.08	0.000	0.000	0.000	0.000	0.000
	Max		298.65	0.021	0.014	-0.014	-0.088	-0.040
	Top 30		104.00	0.007	0.005	-0.005	-0.030	-0.014
Direct estimates from Table 2				0.007	-0.010	-0.005	-0.034	

Appendix L. Kwon and Roth Test of Mediating Mechanism

In this appendix, we explore mechanism tests in the spirit of Kwon and Roth (2026) as an additional check on the commuting-based channel. Their framework is designed to test the sharp null of *full mediation*: The treatment affects the outcome only through the specified mediator(s). Thus, their test is conceptually quite different from our mediation analysis, which is designed to test the null of *no mediation effect*, which permits the existence of alternative paths. Unfortunately, its application is not straightforward in our setting because we allow for multiple mediators, continuous mediator variation, and direct effects of treatment --- we *are* assuming channels other than mediators exist. Nonetheless, we assess whether the Kwon-Roth framework can be informative. To do so, we implement sharp-null mechanism tests after discretizing the two mediators into quantile bins.

Table A17 reports the resulting p-values under alternative discretization schemes. Mediator 1 is the share of commuters using private cars, and Mediator 2 is the share using public transit. The columns labeled "Ordered to match hypothesized effects" assign mediator categories so that larger category values correspond to the direction in which treatment is hypothesized to move each mediator, while the columns labeled "Reverse ordering" reverse the category ordering for the public-transit mediator. We report results using 3, 4, and 5 bins. The rows labeled "Sharp null" test the null that the corresponding mediator, or both mediators jointly, do not account for the treatment effect. The rows labeled "Joint null allowing x% defiers" relax the monotonicity assumption by permitting the indicated share of defiers.

We see the limitations of applying this class of tests in our setting. As Table A17 shows, the conclusions are highly sensitive to the discretization rule, the ordering of mediator categories, and the assumed share of defiers. In some specifications the sharp null is rejected, while in others it is not. We, therefore, do not view these results as sufficiently stable to support formal interpretation and treat this exercise as exploratory only.

Table A17. Exploratory Sharp-Null Mechanism Tests under Alternative Discretizations

	Ordered to match hypothesized effects			Reverse ordering of mediator 2		
Mechanism tested	**3 bins**	**4 bins**	**5 bins**	**3 bins**	**4 bins**	**5 bins**
Sharp null: car-commuting mediator only	0.994	0.811	0.300	0.994	0.811	0.300
Sharp null: public-transit mediator only	0.000	0.000	0.000	0.342	0.841	0.016
Sharp null: both mediators jointly	0.000	0.000	0.000	0.357	0.977	0.415
Joint null allowing 1% defiers	0.421	0.005	0.082	0.721	1.000	1.000
Joint null allowing 2% defiers	0.647	0.633	0.411	0.858	1.000	1.000
Joint null allowing 3% defiers	0.915	0.920	0.962	0.858	1.000	1.000
Joint null allowing 4% defiers	0.569	0.939	1.000	0.858	1.000	1.000
Joint null allowing 5% defiers	0.975	1.000	1.000	0.858	1.000	1.000