VancUber: The Long-Run Effect of Ride-hailing on Public Transportation, Congestion, and Traffic Fatalities*

John Cairncross[†]   Jonathan D. Hall[‡]   Craig Palsson[§]

July 2, 2024

Abstract

We investigate the long-run effect of ride-hailing on public transit ridership, traffic congestion, and traffic fatalities. We estimate the long-run effect by exploiting British Columbia's use of a pre-existing regulation in 2013 to ban ride-hailing from Vancouver. Using difference-in-differences, we show that the estimated effects are sensitive to the choice of control group. Motivated by this, we use the synthetic control method to construct a counterfactual Vancouver. We do not find a statistically significant effect of ride-hailing on our outcomes. To help understand these findings, we conduct a meta-analysis. We find significant heterogeneity in the literature's estimates, but the average estimate is near zero.

KEYWORDS: Ride-hailing, traffic fatalities, traffic congestion, public transportation, meta-analysis

JEL CODES: R40, O33

[†]Department of Economics, University of Toronto, 150 St. George Street, Toronto, Ontario M5S 3G7, Canada, john.cairncross@mail.utoronto.ca

[‡]Department of Economics, Finance, and Legal Studies, University of Alabama, Box 870224, Tuscaloosa, Alabama 35487, USA jonathan.hall@ua.edu

[§]Huntsman School of Business, Utah State University, 3500 Old Main Hill, Logan, Utah 84322, USA, craig.palsson@usu.edu

1 Introduction

New transportation technologies have repeatedly reshaped cities. The invention of steam railways allowed cities to start expanding beyond walking distance (Heblich et al., 2020), the invention of the automobile changed where rich and poor lived within cities (LeRoy and Sonstelie, 1983), and the construction of limited-access highways caused cities to spread out further (Baum-Snow, 2007). Ride-hailing, as exemplified by firms such as Uber, Lyft, and Didi, is the latest new transportation technology to affect cities. The rapid onset of ride-hailing has inspired significant policy debates, and lead to a wave of academic research seeking to measure its impact on cities. Because ride-hailing is so new, most studies have only measured the short-run impact. However, designing effective policy requires understanding the long-run impacts.

We investigate the long-run effect of ride-hailing on three transportation-related outcomes of particular interest to policymakers: (1) public transit ridership, (2) traffic congestion, and (3) traffic fatalities. Each of these outcomes has attracted significant interest from a variety of academic fields, including economics, geography, transportation engineering, epidemiology, and information systems; however, there remains great uncertainty about even the short-run impact, as existing estimates vary in sign and magnitude.

A natural challenge in estimating the long-term effect of ride-hailing on cities is that it is difficult to know what the counterfactual outcome would have been in its absence. Ride-hailing firms spread quickly, with their most popular service (i.e., UberX or Lyft) entering 46 of the 50 largest metropolitan areas in the US and Canada within 22 months, and so it is difficult to find a credible comparison group, especially for estimating long-run effects. For example, in a standard difference-in-differences design, effects estimated in 2015 require comparing cities such as New York City, Los Angeles, and Chicago to cities such as Buffalo, New York; Iowa City, Iowa; and Springfield, Missouri. It is more difficult to accept the parallel trends assumption for cities that are so different. Estimating effects in later years requires even more challenging comparisons.

In this paper, we overcome this challenge by taking advantage of British Columbia's ban on ride-hailing. In 2012, British Columbia used pre-existing regulations to block Uber from entering Vancouver and then successfully excluded

all ride-hailing services until January 2020. Vancouver's unique regulation of ride-hailing provides two important benefits for studying the long-run effects of ride-hailing services. First, the ban lasted a long time in a city comparable to early-entry cities. Second, the long-standing character of British Columbia's taxi regulations gives a strong case for the availability of ride-hailing being exogenous to our measured outcomes.

To estimate what would have happened in Vancouver had Uber or Lyft entered, we take multiple empirical approaches. First, we motivate the analysis using two types of pairwise difference-in-differences comparisons. We compare Vancouver to three natural comparison cities (Seattle, Toronto, and Portland). We next compare Vancouver to all other cities where Uber entered in 2013. We find that the direction and magnitude depend on which comparison cities are chosen.

To avoid the bias of selecting our preferred comparison group, we build a synthetic Vancouver following the synthetic control methodology of Abadie and Gardeazabal (2003) and Abadie et al. (2010). For each of our three outcomes—public transit ridership, traffic congestion, and traffic fatalities—we create a synthetic Vancouver by finding the weighted average of cities in the US and Canada that best approximates Vancouver's time series before the entry of ride-hailing. We estimate ride-hailing's effect for each post-treatment year by comparing the time path of this synthetic Vancouver with the city's actual outcomes. We consider these as our primary estimates.

We find a statistically-insignificant, positive effect of ride-hailing on public transit ridership, traffic congestion, and traffic fatalities. While our estimates are statistically insignificant, Abadie (2020) argues against "the usual practice of conferring point null rejections a higher level of scientific significance than non-rejections" since nonsignificant results are informative, and in some cases, are more informative than statistically significant results. Our confidence intervals, constructed following Firpo and Possebom (2018), are $[-1.3\%, 4.1\%]$, $[-6.1\%, 20.7\%]$, and $[-30.6\%, 198.6\%]$ for transit ridership, congestion, and fatalities, respectively.

To put these estimates in context, we conduct a meta-analysis of the literature estimating the impact of ride-hailing on transit ridership, congestion, and fatalities. After screening 2,597 articles, we end up with a database of 56 relevant articles. We find there is substantial heterogeneity in the literature; for example, estimates for the impact of ride-hailing on public transit ridership range from -38.9% to +146%.

However, we also find that the average estimates on all three outcomes are close to zero and none are statistically significant.[1] Additionally, we find that 25%–41% of articles on each outcome report that they fail to find statistically significant estimates for the impact of ride-hailing. We explore sources of heterogeneity in the estimates, finding that geographic setting matters and some suggestive evidence that empirical methodology matters. We find no evidence of publication bias.

From our study of Vancouver, we conclude that the short- and long-run impact of ride-hailing on transit ridership and traffic congestion is likely to be small for mid-sized cities in the US and Canada. From our meta-analysis, we conclude that the effect of ride-hailing on cities depends on the local context. This is to be expected; there are a variety of mechanisms by which ride-hailing can affect each outcome, and in different contexts, different mechanisms matter more. For example, in cities with small transit agencies, ride-hailing might solve last-mile problems and therefore increase ridership, but in cities with large transit agencies, ride-hailing might not increase the reach of transit and so just be a competitor.[2] However, general statements about the impact of ride-hailing on transit ridership, traffic congestion, and traffic fatalities should reflect that the average estimate is close to zero and not statistically significant. We further note that accepting that the impact of ride-hailing is context-dependent limits the external validity of any given study.

The rest of the paper proceeds as follows. We explain the history of ride-hailing in Vancouver (Section 2), introduce our data (Section 3), and provide a motivating analysis using difference-in-differences (Section 4). We then introduce our preferred methodology of synthetic control (Section 5) and report our results (Section 6) and four robustness tests (Section 7). Section 8 contains our meta-analysis, and we conclude in Section 9.

---

[1]An important exception to this is that, on average, ride-hailing decreases bus ridership while increasing train ridership.

[2]See Hall et al. (2018) for a longer discussion of the mechanisms by which ride-hailing could affect public transit ridership, and see Krishnamurthy and Ngo (2022) and Barrios et al. (2022) for discussions of how ride-hailing could affect congestion and traffic safety, respectively.

2 History of ride-hailing in Vancouver

To identify the long-term effects of ride-hailing on cities, we rely on Vancouver's strict regulation of ride-hailing services. Despite strong latent demand for ride-hailing, the taxi industry successfully blocked these services from entering the Vancouver market. Thus, Vancouver provides a good case study of what a city looks like without ride-hailing services in the long-run.

Uber's early attempt to enter Vancouver in 2012 was nearly successful, but it was ultimately thwarted by British Columbia's Passenger Transportation Board (PTB). Throughout its early years, Uber entered markets following a well-established pattern: offer the service without permission and deal with regulatory concerns later (see Hall et al. (2018) for additional detail on patterns in Uber's entry decisions). Its entrance into Vancouver was supposed to be the same. During the Summer of 2012, Uber provided limited service in Vancouver under its "Secret Uber" program.[3] After exploring the market, Uber planned to launch full service in November 2012. Ahead of the launch, the PTB informed Uber that it would be classified as a limousine company and would therefore have to charge a $75 minimum fee, regardless of trip distance or duration.[4] Uber CEO Travis Kalanick claimed that the company knew of the rule but also that their research found few comparable services were following it; that many limo companies offered rides for less than $75 and that airport limousines had an exemption. Nevertheless, Uber and other ride-hailing services were prevented from entering the market.

Importantly, for this paper, the city's resistance did not come from concerns about public transit, traffic safety, or congestion. The PTB stopped Uber because of concerns about the taxi industry. While taxis opposed Uber in every city it entered, the opposition succeeded to a much greater degree in Vancouver. Its inordinate success preceded Uber. At the beginning of 2012, Vancouver had an abnormally low supply of taxis: 9.4 per 10,000 compared to Montreal's 27 and Toronto's 18.[5] Although the city restricted licenses, it distributed them for a small

---

[3]Uber. 2012. "Regulators Demanding Uber Vancouver Price Increase! We Need Your Help!" November 22, 2012. https://web.archive.org/web/20131023023017/http://blog.uber.com/2012/11/22/helpubervan/

[4]Schelling, Steven. 2012. "Uber town-car service shut down in Vancouver by B.C. Passenger Transportation Board. https://web.archive.org/web/20220119073731/https://www.straight.com/news/uber-town-car-service-shut-down-vancouver-bc-passenger-transportation-board

[5]Brocki, Luke. 2012. "Welcome to Taxiland." The Dependent Magazine, June 1, 2012. https:

fee of 522 Canadian dollars (CAD) per license. License owners could then lease the licenses on a secondary market where access to a full license would sell for 800,000 CAD. Although Vancouver officials acknowledged the severe shortage of taxis, they also recognized the industry's political power and stake in the status quo. Uber recognized their main barrier was the taxi industry. When Vancouver first stopped Uber in November 2012, the company responded with a call to Vancouver residents to contact their representatives and the PTB with the message, "ABOLISH TAXI-PROTECTIONISM - LET UBER CHARGE LESS THAN $75 FOR A RIDE!"[6] Thus preventing Uber's entry protected the pre-existing taxi rents and was unrelated to concerns over the outcomes we are measuring.

Another important feature of Vancouver is that even after the initial failure, Uber wanted to enter the market and there was clearly latent demand for the services. In the Summer of 2014, Uber tried to build more goodwill and brand awareness by delivering ice cream.[7] Then, in October of that year, despite no operations in British Columbia, Uber held a hiring fair in a Vancouver hotel to recruit drivers.[8] While Uber focused on obeying the rules and softening public opinion, other services entered through underground channels. For instance, some Chinese-language companies began offering ride-hailing services illegally.[9] Although such underground services were too small to affect city-wide transportation, they reflect the latent demand for ride-sharing.[10] Over the years, political pressure mounted

//web.archive.org/web/20120707063224/http://thedependent.ca/featured/taxiland/

[6]Uber. 2012. "Regulators Demanding Uber Vancouver Price Increase! We Need Your Help!" November 22, 2012. https://web.archive.org/web/20131023023017/http://blog.uber.com/2012/11/22/helpubervan/

[7]CBC News. 2014. "Taxi app tries for a comeback in Vancouver." July 18, 2014. https://web.archive.org/web/20211021182529/https://www.cbc.ca/news/canada/british-columbia/taxi-app-tries-for-a-comeback-in-vancouver-1.2711847

[8]CBC News. 2014. "Uber Vancouver hiring fair goes on despite moratorium." October 14, 2014. https://web.archive.org/web/20220119091332/https://www.cbc.ca/news/canada/british-columbia/uber-vancouver-hiring-fair-goes-on-despite-moratorium-1.2798663

[9]Quan, Douglas. 2018. "Underground ride-sharing services thrive in B.C. — if you speak Chinese." April 5, 2018. https://nationalpost.com/news/canada/underground-ride-sharing-services-in-b-c-appear-to-be-thriving-amid-crackdown

[10]As an example of how small these underground services were, one reported it was serving 10,000 people per month in the Vancouver area, while in the first month of ride-hailing being legal in Vancouver, there were 45,000 ride-hail rides and 1.5 million taxi rides. Quan, Douglas. 2018. "Underground ride-sharing services thrive in B.C. — if you speak Chinese." April 5, 2018. https://nationalpost.com/news/canada/underground-ride-sharing-services-in-b-c-appear-to-be-thriving-amid-crackdown and Azpiri, Jon. 2021. "Ride-hailing trips in Metro Vancouver outnumber taxis nearly 2-to-1: report." November 12, 2021. https://web.archive.org/web/

to allow ride-hailing platforms to enter the market, and by 2017, the three major political parties announced intent to open British Columbia.[11] In January 2020, Uber entered Vancouver, its last major metropolitan market in North America.[12] The PTB approved Lyft the same week and approved other ride-hailing services in the following months.

Vancouver's regulation provides a unique setting to identify the long-run effect of ride-hailing on cities. Since Uber wanted to enter Vancouver at the same time as other large cities but was rebuffed, it serves as a better comparison city than markets where Uber delayed entry because they were too small.[13] We set Uber's intended entry into Vancouver as 2013. We do this for two reasons. First, the treatment of interest is when cities get UberX, the service that most people associate with Uber, not UberBlack, Uber's first service, which was a high-end limousine service. Most cities did not get UberX until 2013, and we assume that with its early intent to enter the Vancouver market, Uber would have also launched this service in 2013. Second, while Uber's official launch was scheduled for late November 2012, we are using annual data, so it makes more sense to attribute the first year of treatment to 2013. Setting treatment as 2013 means the donor group includes cities like Atlanta, Seattle, and Washington DC. We also conduct a robustness test where we set treatment to 2014 because that is when UberX came to other major Canadian cities like Toronto and Montreal.

3 Data

We gather annual data on public transit ridership, traffic fatalities, and traffic congestion; Uber and Lyft entry dates; and economic indicators for metropolitan areas and cities in the US and Canada.[14] One problem is that the data are reported

---

20230804021148/https://globalnews.ca/news/8370335/metro-vancouver-ride-hailing-report/.

[11]Kelly, Ash. 2017. "NDP stalls on election promise to bring ride hailing to B.C. by end of year." CBC News, October 16, 2017. https://web.archive.org/web/20210724190210/https://www.cbc.ca/news/canada/british-columbia/ndp-stalls-on-election-promise-to-bring-ride-hailing-to-b-c-by-end-of-year-1.4357347

[12]van Hemmen, Michael. 2020. "Vancouver: Uber is Here." January 25, 2020. https://web.archive.org/web/20201020153907/https://www.uber.com/en-CA/newsroom/vancouver-uber-is-here/

[13]Per Hall et al. (2018), Uber entered markets with larger populations first.

[14]In the US, we use Core Based Statistical Areas (CBSA) and in Canada, we use Census Metropolitan Areas (CMA).

at different levels of geographical detail: public transit ridership is reported for transit agencies, traffic congestion is reported for metropolitan areas, and traffic fatalities are reported for cities. For expositional ease, in the analysis we will refer to the unit of observation for all outcomes as a city.[15]

3.1 Public transit ridership

We obtain annual public transit ridership data for 2006 through 2017 from publicly-available reports by the American Public Transit Association (APTA) and TransLink, the primary transit agency for Greater Vancouver. The APTA data feature several Canadian cities, including Toronto, Montreal, and Vancouver, and are agency-mode specific. For each metropolitan area, we use data for the primary transit agency which serves the central city and any commuter rail agencies.

One feature of the APTA data is that each year's report contains data from both the current and previous year. The likely reason for this is that data for some agencies may be unavailable at publication time. Thus the following year's report updates ridership data for the same quarter in the previous year. This revision may also include corrections to misreports in the previous year, for example, because of typographical errors. As a default, we use the later-year report to extract data for the previous year; thus, the 2007 through 2018 reports give us ridership data for 2006 through 2017. Since the universe of reporting agencies is not constant over time, an agency may report in one year and not in the next. We fill any gaps in the data using current-year reports.

The APTA data present a problem with respect to TransLink. TransLink is absent from APTA's Q4 2015 report, and data for TransLink are partial for 2016 and 2017. Even filling the 2014 data using the current-year report leaves a two-year gap

---

[15]We use annual data for three reasons. Most importantly, constructing a synthetic treated unit necessitates choosing the counterfactual time period when Uber would have entered Vancouver in the absence of the pre-existing regulation. This time period then determines the set of potential donor cities used to form the synthetic control. Using annual data, instead of quarterly or monthly data, allows us to have a large donor pool. Moreover, we prefer to use a consistent frequency across specifications, and our congestion data are only available at annual frequency. Finally, using annual data mitigates seasonality concerns. As we are interested in the long-run impacts of ride-hailing, the use of annual data is appropriate in our context. We acknowledge that using annual data means we are measuring the average annual impact, and that we cannot detect potential heterogeneous effects at smaller time scales. For example, a decrease in off-peak congestion that is offset by an increase in peak congestion, as found in Krishnamurthy and Ngo (2022).

for some modes, and total ridership, in the post-period for Vancouver. However, mode-specific ridership data are available from TransLink's annual reports from 2006 through 2017. Accordingly, we use TransLink's data for Vancouver and continue to use APTA data for all other metropolitan areas. Figure A1 compares ridership for TransLink using the two data sources. Correlations between each series are over 90% for all modes except "other," at 76%.

3.2 Traffic congestion

We obtain congestion data for 113 metropolitan areas for 2010 through 2019. This data comes from two sources, TomTom and Inrix. Both services use data from various sources, including smartphones, navigation devices, and roadway sensors to measure travel times. TomTom data from 2012 through 2019 are publicly available. The advantage of the TomTom data is that it entirely contains the treatment period and some of the pre-treatment period. This means the treatment effects can be estimated within the same dataset and we do not have to worry about harmonizing different measures. The challenge, however, is that there is only one pre-treatment year for generating the synthetic control weights. We could still estimate the synthetic control with the one year of pre-treatment data, but to get a better fit with more pre-treatment periods, we need to add more data. For this reason, we supplement the TomTom data using Inrix, which has publicly-available data from 2010 through 2013.

The TomTom and Inrix congestion indices measure how much longer an average trip takes relative to uncongested conditions. That is, an index level of 25 means the average trip takes 25% longer than it would when there is no congestion. We use indices that take the average delay for all daily trips, as measures focusing on specific times of day (such as the morning peak) are not available over our entire sample period.

The TomTom and Inrix measures are highly correlated across metros within overlapping years (the correlation is greater than 80% in 2012 and 2013). However, the correlation between changes in the TomTom and Inrix indices across CBSAs within overlapping years is low. For example, the across-city correlation of the change in the respective congestion measures from 2012 to 2013 is only about 0.1. However, since we treat all cities similarly when combining data sources, the

8

synthetic control methodology should generate weights that produce an unbiased estimate of treatment effects. We combine the TomTom and Inrix data by taking the average of the two measures in the two overlapping years.

3.3   Traffic fatalities

We obtain traffic fatality data for the primary city of each metropolitan area for 2004 through 2017 for 120 cities.[16] Our US fatalities data come from the Federal Accident Reporting System (FARS). We aggregate this data to the city level. Canadian fatalities data were collected from publicly-accessible websites maintained by provincial and municipal governments.

3.4   Uber and Lyft entry

We extend the ride-hailing entry data from Hall et al. (2018) by extending our Uber entry data through 2020 and adding Lyft entry dates. We use official press releases where possible, but also use newspaper articles, blog posts, and social media posts. We corroborated our Lyft entry dates with data collected for Teltser et al. (2021). For Uber, we consider only UberX (not, e.g., UberBlack). Where several municipalities within a single metropolitan area received ride-hailing in different years, we take the principal city as providing the metropolitan entry date. We also exclude any metros where ride-hailing entered, only to subsequently leave either temporarily or permanently.

3.5   Population, gas prices and unemployment

We use data on metropolitan populations to normalize public transit ridership and fatalities. We collect data on population for US and Canadian metropolitan areas from the US Census Bureau and Statistics Canada. Additionally, in a robustness check, we add gas prices and metropolitan unemployment rates to the list of covariates we use to construct a synthetic "Vancouver." For the US, these data

---

[16]Two exceptions to only including the primary city are that we include both Minneapolis and St. Paul as well as Dallas and Ft. Worth, as they are the only metropolitan areas with two major cities at their center.

come from the US Energy Administration and the Bureau of Labor Statistics, respectively. Canadian data for both measures are from Statistics Canada.

4   Motivating analysis

We start by looking at the trends for our three outcomes in Vancouver and three natural comparison cities: Seattle, Toronto, and Portland, Oregon. The three cities are natural comparisons because of their similarities to Vancouver with size and public transit systems, but they are also convenient because Uber entered the three cities in different years. The trends are plotted in Figure 1, with the year of Uber entry marked with a vertical line (Seattle, 2013; Toronto, 2014; Portland, 2015).

With the three comparison cities, one could try to estimate the effect of ride-hailing on each outcome using a simplified difference-in-differences exercise; however, the results are sensitive to the reference city. Public transit ridership provides a good example. Comparing the changes in the growth rate of per capita transit ridership in Vancouver to those in Seattle, Toronto, and Portland before and after the entry of ride-hailing yields simple difference-in-differences estimates of the effect of ride-hailing on ridership, reported in Table A1, of 4.1%, 1.8%, and -3.8%, respectively. The results for congestion and fatalities are similarly sensitive to the city chosen.

To further emphasize estimates' sensitivity to the choice of comparison city, we conduct a second difference-in-difference exercise using all cities that ride-hailing entered between 2013 and 2015. For each outcome-comparison city, we run a pairwise difference-in-differences regression using observations from that city and Vancouver, after removing pre-period trends.[17] We plot the resulting estimates in frequency histograms, in Figure 2. We explicitly mark estimates for Portland, OR, Seattle, and Toronto, although due to detrending these estimates are not the same as those reported in Table A1.

---

[17]More specifically, we take the natural log of each outcome and evaluate changes relative to the last pre-treatment year for each city. This accords with our main analysis, detailed in Section 5. We set Vancouver's "treatment" year as 2013. We remove pre-period trends for cities in each treatment year (2013, 2014 and 2015) by regressing pre-treatment outcomes on city-specific time trends and predicting residuals for all years. We then run a series of pairwise regressions for each treated city, retaining observations for that city and Vancouver, in which the right-hand side variables are year and city fixed effects, plus an indicator for the presence of ride-hailing. The coefficient on this indicator variable is our estimated treatment effect.

In Figure 2, it is easy to see that difference-in-difference estimates of ride-hailing's effects are sensitive to the choice of comparison city. Although for all three outcomes estimates cluster around zero, their dispersion, particularly for congestion and fatalities, suggests the need for formalizing the choice of the comparison group. This is a major advantage of the synthetic control methodology, to which we now turn.

5 Methodology

5.1 Synthetic treatment

To estimate the long-term effect of ride-hailing on cities, we use a modified synthetic control analysis. Synthetic control methods provide some key advantages for the challenges we face. First, our analysis relies on one untreated unit, which traditionally would make it a comparative case study. But the synthetic control method of Abadie et al. (2010) formalizes the selection of comparison groups in comparative case studies and provides methods for statistical inference. Specifically, in studies with panel data and a small number of treated units, the methodology constructs a synthetic comparison group as a weighted average of untreated units. The treatment effect in period t is then the difference in the outcome variable between the treated unit and its synthetic counterpart. See Appendix A for a formal description of the synthetic control method and additional details on inference, and see Abadie (2021) for a longer review.

Second, synthetic control has several advantages in our situation. Crucially, synthetic control relaxes the parallel trends assumption (Abadie, 2021), which may be problematic in our context. Moreover, Goodman-Bacon (2021) shows that two-way-fixed-effect estimation may be problematic in contexts where treatment effects are dynamic, as seems reasonable here. Our aim is to estimate the dynamic effects of ride-hailing entry over the long term for Vancouver, an objective for which synthetic control is well-suited.

We adapt synthetic control to create a synthetic *treated* Vancouver. Given a treatment date, our donor group is the set of all cities where Uber or Lyft entered in that year. As explained in Section 2, we choose 2013 as our baseline date. The donor group is then weighted to construct a synthetic treated Vancouver that

best matches the pre-treatment level of the variables we are matching on. The number of observables on which to match is potentially large, and there is no consensus on how best to choose them (see the discussion in Abadie (2021)). We match on the pre-treatment average of the outcome variable of interest in the pre-treatment period, following Abadie and Gardeazabal (2003) and Abadie et al. (2010), plus the level of the outcome variable in even years before 2012. In our baseline specifications we exclude 2008 and 2009 — the Great Recession — from both the average pre-treatment outcome and the set of individual years used to match. We include these in a robustness check. In a further robustness check, we add average pre-treatment unemployment and gas prices to the list of variables we match on, as changes in gas prices and unemployment may predict changes in our outcomes. This changes the set of donors and their weights for both ridership and fatalities, although not for congestion.

It is important to note how synthetic control works across different outcomes. Given a set of donor cities, the objective of synthetic control is to select weights that minimize the difference between the synthetic outcome and the treated outcome. Since the data generating process is different for each outcome, the donor weights will likewise be different. For example, putting a non-zero weight on Seattle may minimize the mean squared prediction error for public transit, but when analyzing congestion it may be optimal to give Seattle a weight of zero. Thus, as is seen in the Appendix, while the donor groups for each outcome are roughly the same size, the number of cities that receive non-zero weights and their identities differ for each analysis.

Our outcome of interest is the growth rate in the three transportation variables: public transit ridership, congestion, and traffic fatalities. Before calculating growth rates for transit ridership and traffic fatalities, we normalize the variables by population. Then, for all outcomes, we take the log difference between that year and our base year, 2012, the last pre-treatment year in our main specification.[18] Thus, our synthetic treatment procedure finds the weighted average of cities whose

---

[18]Our rationale for subtracting the log 2012 level of the outcome variable is twofold. First, Vancouver has the highest per-capita transit ridership of any of our potential donors, and a synthetic Vancouver using this measure could not be attained as a convex combination of those cities. Second, our donor cities are very heterogeneous. Because we are interested in changes in our outcomes, using levels of cities' outcome variables to weight the donor group would be inappropriate.

growth relative to 2012 best matches Vancouver, and uses this weighted average to predict how Vancouver would have grown relative to 2012 had ride-hailing entered in 2013.

5.2 Inference

Abadie et al. (2010) and Abadie et al. (2015) propose permutation-based inference, in which statistical significance is inferred by comparing the estimated treatment effect for the unit that was treated, with similar placebo effects for untreated units constructed by running synthetic control on those units. We follow Abadie et al. (2010) and Abadie et al. (2015) in performing a test of overall significance by comparing the post-treatment root mean squared prediction error (RMSPE) for Vancouver with the RMSPE for placebo cities. If the null hypothesis of no treatment effect holds, we expect the post-treatment RMSPE for the placebo cities to be larger than Vancouver's. *p*-values are "standardized" by scaling post-treatment RMSPEs by pre-treatment RMSPEs. This ensures that we do not identify a statistically significant overall effect in Vancouver just because its pre-treatment fit tends to be poor relative to placebo cities.

Given the wide variety of estimates in the literature, we consider it important to be able to rule out certain effects with a relatively high degree of confidence. Firpo and Possebom (2018) propose inverting permutation tests to obtain confidence intervals.[19] We also implement this procedure. Note that we invert overall *p*-values that account for the quality of pre-treatment fit, ensuring a consistent interpretation of *p*-values and confidence intervals.

6   Results

6.1   Public transit ridership

We first estimate the synthetic control for public transit ridership. The analysis puts non-zero weights on nearly every donor city, as seen in Table A7, but two cities receive disproportionate weights: San Diego (0.43) and Oklahoma City (0.20).

---

[19] Note that in Firpo and Possebom (2018)'s procedure, confidence is expressed as a fraction, where the numerator must be an integer and the denominator is given by the sum of treated and placebo units. We choose confidence levels as close as possible to 95%.

When looking at these weights, some might question whether these two cities are natural comparisons for Vancouver. But this is the strength of the synthetic control analysis. We have already shown that picking a "natural" comparison city results in different estimates of the treatment effect depending on the city. With the synthetic control analysis, the weights are objectively determined by the donor cities that most closely approximate the data-generating process of the Vancouver public transit system. By presenting all estimates, we hope to show that the overall story is consistent regardless of the assumptions behind choosing comparison cities.

Figure 3, panel (a) shows the estimated treatment effects for Vancouver and all placebos. The pre-treatment "effects" are close to zero, showing that Vancouver matches our synthetic Vancouver well; furthermore, Vancouver's match to our synthetic Vancouver is good relative to that of the placebo cities to their synthetic versions. Table 1 reports the point estimates, standardized p-values, and the confidence interval for the average effect. Our point estimates for the effect of ride-hailing on public transit ridership are positive for the first three years and negative for the last two.[20] None of these point estimates are statistically significant. Our p-value for the overall RMSPE likewise indicates that we cannot reject that the observed effects occurred simply by chance.

Another piece of evidence for a small or null effect is that the point estimates do not show a dose-response relationship. It should not be surprising to find ride-hailing had no effect in its early days because its market share was small. But since ride-hailing ridership has grown over time, we would expect the early effects to be smaller than the later effects.[21] The point estimates, however, show

---

[20]Part of the large effect in 2017 is likely due to Vancouver extending its Sky Train by 11-kilometers in December of 2016. It is difficult to separate out the effect of the extension on total ridership, as the extension likely increased ridership on other lines and modes by increasing the reach of the transit system, but also replaced a popular bus route, which accounts for a third of the weekday ridership on the extension in 2017 (Talmazan, Yuliya. 2017. "Evergreen Line ridership reaches 30,000 trips a day in January." Global News, March 1, 2017. https://web.archive.org/web/20181127045347/https://globalnews.ca/news/3280798/evergreen-line-ridership-reaches-30000-trips-a-day-in-january/). However, if these effects cancel out, then we can approximate the change in transit ridership in the absence of the extension by what happened on all other lines and modes. Doing so implies that total transit ridership in the absence of the extension would have increased by 3.7% (rather than 5.7%). This suggests that had Vancouver not extended the Sky Train we would have found a treatment effect in 2017 of -0.063.

[21]For evidence that ride-hailing ridership has grown with time (pre-pandemic), see Hall et al. (2018) or Dean (2021).

insignificant positive effects early and insignificant negative effects later on. This could be evidence that the marginal user is changing over time. Early users might have been using ride-sharing to solve last-mile problems with public transit, thus increasing ridership, while later users were using ride-sharing as a substitute for public transit, thus decreasing ridership. But it would be convenient for those countervailing effects to nearly exactly cancel each other out. The more likely explanation is that ride-sharing's effect is relatively small, and the fluctuations are part of normal variation in ridership.

Furthermore, Table 1 reports a confidence interval for the effect of ride-hailing on log transit ridership of $[-0.013, 0.040]$, which translates to $[-1.3\%, 4.1\%]$.[22] These results rule out large effects in Vancouver.

6.2 Traffic Congestion

Next, we perform the synthetic control analysis for traffic congestion. Similar to transit ridership, most donor cities receive non-zero weights, but the weights are different. For example, the two cities with the greatest weight in this analysis are Denver (0.54) and Boston (0.36). The weights are different because the synthetic control attributes weights according to the outcome variable's data-generating process. Since this process differs for transit ridership and traffic congestion, the matched cities are different. Again, this is an advantage of the synthetic control analysis because it does not require us to assume that the correct counterfactual cities for one outcome are the right ones for another outcome.

Figure 3, panel (b) and Table 1 report estimated treatment effects for Vancouver, using the combined Inrix and TomTom dataset described above. Again, the pre-treatment "effects" are close to zero, showing that Vancouver matches our synthetic Vancouver well, and, Vancouver's match to our synthetic Vancouver is good relative to that of the placebo cities to their synthetic versions. Our point estimates are all positive, and range from 0.2% to 12.5%. Point estimates in all post-treatment years are insignificant, and our p-value for the overall effect indicates that we cannot reject that the observed effects happened by chance in any post-treatment year. The 95% confidence interval for the average effect is $[-6.1\%, 20.7\%]$.

---

[22]This conversion is made by exponentiating the ends of the confidence interval and subtracting one.

Of the three outcomes, traffic congestion is the most likely to exhibit a dose-response relationship, yet we find no evidence for it. If ride-sharing did increase congestion, the main mechanism would be that there are more cars on the road. Thus, as ride-sharing becomes more prevalent, and more ride-sharing cars are on the road, the effect should be more apparent. Yet, in Figure 3, the effect of ride-hailing on congestion does not grow over time. This gives greater evidence that if ride-sharing has any effect on congestion, it is small.

As with public transit ridership, the pattern in the point estimates displayed in Figure 3 provides further evidence that the effect is small, as the effect of ride-hailing on congestion fails to grow with time, even though ride-hailing ridership grows with time.

6.3   Traffic Fatalities

Finally, we do the synthetic control for traffic fatalities. This one differs from the previous two outcomes because it puts all of the weight on only two cities: St. Paul (0.84) and Boston (0.16). This exclusive focus on two cities may contribute to the imprecise estimates we find below.

Figure 3, panel (c) and Table 1 show the estimated treatment effect for city-level traffic fatalities, normalized by CBSA population. In this case, the pre-treatment "effects" are not as close to zero as those of transit ridership and traffic fatalities. The main deviation is during the Great Recession, years we are not trying to match. As before, Vancouver's match to our synthetic Vancouver is good relative to that of the placebo cities to their synthetic versions. We again observe statistically insignificant effects in all post-treatment years. However, in contrast to transit ridership and traffic fatalities, we do observe a dose-response relationship for the impact of ride-hailing on traffic fatalities as our point estimates grow throughout post-treatment to a high of 92% in 2017. However, given that Uber and Lyft accounted for 5.6% of vehicle miles traveled in the core counties of six large US cities in September, 2018, (Balding et al., 2019), this requires that ride-hailing has more than fifteen times the crash risk per mile as other vehicle travel. Uber's own analysis finds the rate of fatal crashes of Uber vehicles per vehicle mile traveled is half that of the US average (Uber Technologies, 2019). This is, in part, due to ride-hailing trips being different than the average automobile trip, as they are more

urban and at slower speeds. Furthermore, ride-hailing can affect crash rates by changing who is driving and under what circumstances (for example, by changing the amount of drunk driving). Given this, it is implausible that ride-hailing is doubling the number of fatal crashes. Our confidence interval for the average effect is very large — between -30.6% and +198.6%. One possible reason that we observe effects of this magnitude is that, relative to our other outcomes, pre-treatment fit for fatalities is generally poor. On average across cities, the pre-treatment RMSPE is 0.161 for fatalities, relative to 0.028 for ridership and 0.049 for congestion. Our confidence interval is so large as to be uninformative about the effect of ride-hailing on traffic fatalities.

7   Robustness

We perform four robustness checks with the goal of understanding the sensitivity of our results to changes in the set of predictors, potential donors, and donor weights. These are reported in Table 2, where the first column reports our baseline results for comparison. In designing our robustness checks, we follow recommendations by Abadie (2021). Our first check is to add average pre-period CBSA gas prices and unemployment rates to the set of variables we match on to create a weighted average of donor cities. Gas prices and unemployment rates are known predictors of whether to travel and which travel mode to use (Taylor and Fink, 2013). Additionally, British Columbia introduced a carbon tax in 2008 which steady increased until 2012; suggesting gas prices may be a particularly important predictor. As can be seen in Column (2), adding these additional predictors has no effect on the statistical significance of the estimated effects of ride-hailing on public transit ridership, traffic congestion, and traffic fatalities. For transit ridership, adding average gas prices and unemployment increases the point estimate slightly, but the estimate remains statistically insignificant. For congestion, adding gas prices and unemployment to the set of predictors has, surprisingly, no effect on the allocation of weights to donors, and therefore no effect on the results. For traffic fatalities, we find a positive point estimate of similar magnitude to our baseline. The estimated effect remains insignificant.

Our second robustness check revises our baseline to include the Great Recession years of 2008 and 2009 in the set of pre-treatment matching variables, plus the

additional predictors described above. As our data for congestion start in 2010, we perform this robustness check for ridership and fatalities only. As can be seen in Column (3), adding the Great Recession years while retaining alternate matching has no effect on the statistical significance of overall estimates for either ridership or traffic fatalities. The ridership estimate is revised slightly downward, and is slightly more precisely estimated than Column (2). The fatalities estimate is revised upward, but remains highly imprecise.

Our third robustness check is a leave-one-out analysis. For each outcome, we re-run the baseline analysis alternately omitting each of the donors that were assigned positive weight in the baseline synthetic control analysis. We rerun the synthetic control analysis for each donor, generating new point estimates and confidence intervals. This checks the robustness of our results to changes in the set of potential (and actual) donors. In Column (4), we report the average point estimate and confidence intervals over all leave-one-out synthetic control analyses. On average, we find results that are comparable to our baseline. For ridership and congestion, the average point estimates are very close to our baseline, as are average confidence intervals. For fatalities, we find that the average leave-one-out point estimate is slightly larger than our baseline and the average confidence interval is somewhat larger.

Our fourth robustness check is in the spirit of Abadie et al. (2015), who backdate German reunification as an "in-time placebo" test. Unique to our analysis is that the time of "treatment" — non-entry of Uber into Vancouver — is not precisely defined. While our choice of 2013 is reasonable given that Uber attempted to enter Vancouver with UberBlack in November 2012 (and, conditional on success, would have likely launched UberX in Vancouver in 2013), it is also true that ride-hailing did not enter other Canadian cities until 2014 (perhaps in response to Uber's failed attempt to enter Vancouver). We therefore perform robustness with respect to our treatment date by changing the year of hypothetical ride-hailing non-entry to 2014. We then take as potential donors the set of cities where ride-hailing actually entered in that year. Choosing 2014 as an alternate treatment date changes the donor pool entirely, and includes other Canadian cities such as Toronto. In keeping with our baseline, we continue to normalize outcomes by their 2012 level, and use exactly the same set of predictors as in the baseline (average outcome up to and including 2012, and alternating years before, excluding Great Recession years).

Column (5) reports the results. In terms of overall effects, with 2014 "treatment" we estimate a negative effect on transit ridership. This in contrast to our baseline. The signs on estimated effects for traffic congestion and fatalities, however, continue to have the same sign as previously. More importantly, all estimated effects continue to be insignificant, and we estimate confidence intervals for all three outcomes that are larger than in our baseline. Moreover, in Figure A3 we plot the time path of estimated dynamic effects with 2014 treatment. These are broadly similar to our baseline. Overall, the results of this robustness check reinforce the message that the effect of ride-hailing on transit ridership, traffic congestion, and traffic fatalities is likely to be small.

In addition to our robustness checks, we test whether the effect of ride-hailing on public transit ridership differs by mode. As Figure A4 and Table A2 show, we do not find statistically significant effects on either bus or rail.[23]

8 Meta-analysis

Our main results and robustness checks consistently show that ride-hailing's effects on congestion and public transit ridership are small and statistically insignificant. However, reports in the news, testimonies to legislatures, and even research in peer-reviewed journals claim that ride-hailing's effects are large and significant.[24] To investigate whether our results are anomalous, we conduct a meta-analysis. We follow the guidelines in Havránek et al. (2020) and Borenstein et al. (2021).

---

[23]There are transit modes that are not bus or rail, notably ferry in Vancouver and paratransit services everywhere. Given the varied nature of services in this category, it is not reasonable to compare them.

[24]For examples, see Fitzsimmons, Emma and Winnie Hu. 2017. "The Downside of Ride-Hailing: More New York City Gridlock." The New York Times. March 6, 2017. https://web.archive.org/web/20240418002620/https://www.nytimes.com/2017/03/06/nyregion/uber-ride-hailing-new-york-transportation.html; Marx, Paris. 2019. " Uber is convenient for city commuters—but bad for cities." NBC News. April 25, 2019. https://web.archive.org/web/20230608163102/https://www.nbcnews.com/think/opinion/uber-convenient-city-commuters-bad-cities-ncna995626; City Council, City of New York. 2018. "Transcript of the Minutes of the Committee on For-Hire Vehicles. April 30, 2018. https://legistar.council.nyc.gov/View.ashx?M=F&ID=6283895&GUID=22FC3C58-BDDE-4411-8962-59436506426D; the submissions and oral presentations to the Select Standing Committee on Crown Corporations of the Legislative Assembly of British Columbia on January 8, 2018, https://www.leg.bc.ca/parliamentary-business/committees/41stParliament-2ndSession-cc/meetingdocuments; and the many articles cited below.

Appendix C provides more details on how we conduct the meta-analysis.

We found 56 papers that satisfied the criteria for inclusion in the meta-analysis. These articles are summarized in Tables 3–8. These tables report the main effect and include any heterogeneous treatment effects mentioned in the article's title, abstract, introduction, or conclusion. Descriptive statistics about the papers are reported in Table 9. One-third of the articles estimate that ride-hailing does not have a statistically significant impact on our outcomes of interest. Only 20% of papers estimate a long-run effect; and the literature is heavily focused on the United States, with 70% of articles based on US data. 68% of the literature use some panel data method, with 43% specifically using difference-in-differences.[25] The remaining 32% of the articles do not use any panel data methods, instead using time-series methods, survey data, or traffic demand models to estimate the counterfactual outcomes in the absence of ride-hailing.

We compare these results in the literature to our results in Vancouver in two ways. First, we look at the distribution of effect sizes in the literature, reported in Table 10. Second, we use a random-effects meta-analysis, where the weighted average is calculated by giving more weight to more precise estimates; that is, effect sizes are weighted by the inverse sum of the estimated variance of the estimate and the variance between studies. We recognize that different disciplines may find some articles' methodologies more convincing than others. For the sake of this analysis, we choose to remain agnostic about the relative credibility of different methodologies.

While the literature contains a wide range of estimates for the impact of ride-hailing on public transit ridership, our null result is consistent with the *average* result in the literature. Table 10 shows that estimates of the average effect of ride-hailing on public transit ridership range from -38.9% to +146.0%; however, the median estimate is -0.01% and when the results are weighted in the meta-analysis in Table 11, the mean effect size across all studies of -1.7%, neither of which are statistically significant. Our results are consistent with the 9 (of 26) papers that find no statistically significant effects, while 59% of existing estimates are large enough to lay outside our estimated confidence interval.

Breaking down the effect of ride-hailing on public transit ridership by mode,

---

[25] As Tables 3–8 reports, many articles use multiple methods, however, for the sake of Table 9, we have made these categories mutually exclusive.

we find that the literature estimates that ride-hailing decreases bus ridership while increasing rail ridership. Table 10 shows that 75% of estimates for bus and rail have the same sign (negative for bus, positive for rail), and Table 11 finds ride-hailing decreases bus ridership by 4% and increases rail ridership by 2%, both of which are statistically significant.[26] These estimates share the same sign as our mode-specific estimates reported in Figure A4 and Table A2.

Turning to the impact of ride-hailing on traffic congestion, we again find a wide range of estimates, with our null estimate again consistent with the average result. Table 10 shows that estimates of the average effect of ride-hailing on traffic congestion range from -4.5% to +29.0% for travel time and -0.9% to 83.5% for VKT, both of which are notably smaller ranges than the estimates for the impact on public transit ridership. The median estimates are +0.2% and +1.5%, respectively, and Table 11 reports that the mean effect size in the meta-analysis is -0.9% and -0.4%, respectively. These estimates for the average effect are not statistically significant, consistent with our estimated null result. 7 of the 28 other papers estimating the impact of ride-hailing on traffic congestion also find a null effect. We find that 14% of existing estimates are large enough to lay outside of our confidence interval.

Our estimates for the impact of ride-hailing on traffic fatalities are too imprecise to be informative; however, we can still summarize the findings from the literature. Table 10 reports that estimates range from -25.3% to +3.3%, with a median of -0.2%, while Table 11 reports an average effect size of -4.1%. Neither of these estimates of the average effect are statistically significant, and 5 of the 12 papers in the literature do not find a statistically significant effect.

8.1 Exploring the heterogeneity in estimates

While our results are consistent with the average results in the literature, the ranges in Table 10 show there is significant heterogeneity in the estimates. Such heterogeneity could come from several sources: publication status, study location, empirical strategy, etc. To learn what sources matter the most, we analyze how the effect sizes vary by sources of heterogeneity using a random-effects

---

[26]Tests for whether the median estimates for bus and rail are different from zero are also statistically significant at the p = 0.006 and 0.022 levels.

meta-regression.[27] We regress the point estimates on various possible sources of heterogeneity. Each estimate is weighted by the inverse of the sum of the (estimated) variance of the estimate and the variance between studies. We include, but do not report, fixed effects for each outcome type (e.g., "bus ridership," "travel times", "fatalities"). Table 12 reports the results. The first column reports results for all outcomes, while the remaining columns report results independently by transit ridership, traffic congestion, and fatalities.

Our most consistent finding from Table 12 is that there are large and statistically significant differences in estimates between studies set in the United States, and those set in European countries.[28,29] This finding that context matters is consistent with other articles' reports of significant heterogeneity in the impact of ride-hailing within the same country, such as Hall et al. (2018), who find that Uber increases transit ridership "more in larger cities and for smaller transit agencies," and Lee et al. (2022) and Li et al. (2022) who find large differences between more and less compact cities.[30]

Additionally, there is evidence that the empirical methodology used matters. For estimates of the impact of ride-hailing on traffic congestion, these differences are statistically significant. However, for estimates of the impact on fatalities, this analysis is complicated by the fact that being set in a European country is perfectly colinear with using a panel data method other than differences-in-differences.

---

[27]Appendix Table A3, A4, and A5 also report the distribution of estimates in various subgroups, providing an alternative way of checking of heterogeneity. However, sample sizes are often very small, which is why we choose to pool the estimates and use a regression to test for heterogeneity while adjusting for differences between outcome types.

[28]There are no studies set in Canada that report standard errors.

[29]Our finding that the geographic setting matters for the impact of ride-hailing raises the question of how our estimates for Vancouver compare to studies set in the US and Canada, rather than those from the entire world. Table A6 reports the meta-analysis estimates of the average treatment effect when the sample is restricted to studies set in the US (there are no studies we can include set in Canada). We find that the estimate of the average treatment effect almost never changes by more than one percentage point, with an import exception for traffic fatalities. This appears to contradict our findings in Table 12 that the geographic setting matters, but results from 70% of studies in our sample being set in the United States.

[30]Taking account of these heterogeneous treatment effects allows us to make more specific comparisons between our estimates and the most relevant estimates from the literature. Our estimates for the impact of ride-hailing on public transit ridership match the estimates from Hall et al. (2018) for a large city with high public transit ridership, but our estimates for the impact on transit ridership and traffic congestion do not match those from Lee et al. (2022) and Li et al. (2022), which both predict reduced transit ridership and increased traffic congestion for more compact cities.

Thus, the -18% change attributed to estimates of the impact of ride-hailing on fatalities due to being set in a European country is also in whole or in part due to using a panel data method other than difference-in-differences. The difference in estimates associated with using difference-in-difference versus a different panel data method is likely because the majority of studies using other panel data methods do not adjust for overall time trends affecting all areas, regardless of the presence of ride-hailing, and so attribute these trends to the ride-hailing.

In contrast to our findings for Vancouver, we find that estimates of the long-run effect are more extreme than the estimates of the average effect. We are concerned that many of these findings are due to the composition of the control group that this long-run effect is being estimated based on. By the time many studies are measuring the long-run effect they are comparing large cities to significantly smaller cities. As mentioned in the introduction, in a standard difference-in-differences design, effects estimated in 2015 require comparing cities such as New York City, Los Angeles, and Chicago to cities such as Buffalo, New York; Iowa City, Iowa; and Springfield, Missouri.

8.2 Publication bias

A final concern is publication bias. While our finding of small, statistically insignificant results seems to go against the received wisdom, it could be that many studies have found similar results but only those that have large, statistically significant results get published. We look for publication bias using two methods. First, we test for an asymmetric funnel plot. Second, we compare published and unpublished papers on whether they report statistically significant estimates and whether they explore heterogeneous treatment effects. The details of these tests, plus further discussion, are included in Appendix D, but we conclude there is little evidence for publication bias. The most surprising finding is that null results are common.

9   Conclusion

A better understanding of the long-run effects of ride-hailing services, such as Uber, is necessary for policymakers to respond optimally to their entry. Unfortunately,

analyzing the long-term effects has been difficult because ride-hailing spread across cities quickly, making it difficult to find a credible comparison group. We address this challenge by using British Columbia's ban on ride-hailing services, which prohibited such services from entering Vancouver until February 2020.

We began our analysis using difference-in-differences, finding that the direction and magnitude of the treatment effect depends on which comparison cities we use. This motivates us to use synthetic control in order to avoid the bias from our decision of the comparison city. Doing so, we find little evidence that ride-hailing would have had large effects on public transit ridership or traffic congestion in Vancouver, even after six years. Our estimates for the effect of ride-hailing on traffic fatalities are too imprecise to be informative.

To better understand these null results, we conduct a meta-analysis of the related literature. We find substantial heterogeneity in estimated treatment effects; however, the average estimates are close to zero and not statistically significant. Geographic setting is an important source of heterogeneity in the estimates. We find no evidence for publication bias.

Our estimates for Vancouver are informative for the long-run impact of ride-hailing on transit ridership and congestion for other mid-sized US and Canadian cities, suggesting that the impact of ride-hailing for these cities are likely small. In combination with our meta-analysis, we conclude that the effect of ride-hailing depends on the local context. Unfortunately, this limits the external validity of any given study. However, general statements about the impact of ride-hailing on transit ridership, traffic congestion, and traffic fatalities should reflect that the average estimate is small and statistically insignificant.[31]

It is also possible that the long-run effects of ride-hailing have yet to manifest themselves. As ride-hailing grows in popularity and as it begins affecting long-term choices such as vehicle ownership and land use (e.g. Gorback, 2020), its effect on cities may change.

Yet another possibility is that we do not find a longer-term effect because ride-hailing helps or hurts these outcomes through different mechanisms, and these effects are canceling out. For example, ride-hailing can help public transit

---

[31]An alternative interpretation of our meta-analysis is to assume there is one true effect, and that the heterogeneity in estimates is just noise. In this case, the conclusion would be that the true effect on all three outcomes is zero. Given the findings within and across studies of substantial heterogeneity based on local context, we do not believe this is true.

by making it easier to get to and from train stations but hurt transit by being an alternative. Likewise, ride-hailing may worsen congestion in specific areas or times while helping in others. Policies that address specific mechanisms by which ride-hailing causes socially harmful effects will be helpful, such as subsidizing ride-hail trips that connect with transit (Agrawal and Zhao, 2023) or congestion pricing (e.g., Hall, 2021, Herzog, 2021).

References

Abadie, Alberto (2020) "Statistical Nonsignificance in Empirical Economics," *American Economic Review: Insights*, Vol. 2, No. 2, pp. 193–208, DOI: 10.1257/aeri.20190252.

———— (2021) "Using Synthetic Controls: Feasibility, Data Requirements, and Methodological Aspects," *Journal of Economic Literature*, Vol. 59, No. 2, pp. 391–425, DOI: 10.1257/jel.20191450.

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller (2010) "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program," *Journal of the American Statistical Association*, Vol. 105, No. 490, pp. 493–505, DOI: 10.1198/jasa.2009.ap08746.

———— (2015) "Comparative Politics and the Synthetic Control Method," *American Journal of Political Science*, Vol. 59, No. 2, pp. 495–510, DOI: 10.1111/ajps.12116.

Abadie, Alberto and Javier Gardeazabal (2003) "The Economic Costs of Conflict: A Case Study of the Basque Country," *American Economic Review*, Vol. 93, No. 1, pp. 113–132, DOI: 10.1257/000282803321455188.

Agarwal, Saharsh, Deepa Mani, and Rahul Telang (2023) "The Impact of Ride-Hailing Services on Congestion: Evidence from Indian Cities," *Manufacturing & Service Operations Management*, Vol. 25, No. 3, pp. 862–883, DOI: 10.1287/msom.2022.1158.

Agrawal, David R. and Weihua Zhao (2023) "Taxing Uber," *Journal of Public Economics*, Vol. 221, p. 104862, DOI: 10.1016/j.jpubeco.2023.104862.

Alexander, Lauren P. and Marta C. González (2015) "Assessing the Impact of Real-time Ridesharing on Urban Traffic using Mobile Phone Data," in *UrbComp'15*, URL: http://humnetlab.berkeley.edu/wp-content/uploads/2020/02/Real-time-Ridesharing_Alexander.pdf.

Anderson, Michael L. and Lucas W. Davis (2021) "Uber and Alcohol-Related Traffic Fatalities," Working Paper 29071, NBER.

Babar, Yash and Gordon Burtch (2020) "Examining the Heterogeneous Impact of Ride-Hailing Services on Public Transit Use," *Information Systems Research*, Vol. 31, No. 3, pp. 820–834, DOI: 10.1287/isre.2019.0917.

Balding, Melissa, Teresa Whinery, Eleanor Leshner, and Eric Womeldorff (2019) "Estimated TNC Share of VMT in Six US Metropolitan Regions,"Technical report, Fehr & Peers.

Barreto, Yuri, Raul Silveira Neto, and Luís Carazza (2021) "Uber and Traffic Safety: Evidence from Brazilian Cities," *Journal of Urban Economics*, Vol. 123, p. 103347, DOI: 10.1016/j.jue.2021.103347.

Barrios, John M., Yael V. Hochberg, and Hanyi Yi (2022) "The Cost of Convenience: Ridehailing and Traffic Fatalities," *Journal of Operations Management*, Vol. 69, No. 5, pp. 823–855, DOI: 10.1002/joom.1221.

Basak, Ecem, Mary Beth Watson-Manheim, and Ali R. Tafti (2020) "Exploring the Causal Mediation Effects of Public Transit Ridership on the Relationship between Ride-Sharing Services and Traffic Congestion: An Empirical Investigation of Uberx in the United States," working paper, SSRN.

Baum-Snow, Nathaniel (2007) "Did Highways Cause Suburbanization?" *Quarterly Journal of Economics*, Vol. 122, No. 2, pp. 775–805, DOI: 10.1162/qjec.122.2.775.

Bekka, Alina, Nicolas Louvet, and François Adoue (2020) "Impact of a Ridesourcing Service on Car Ownership and Resulting Effects on Vehicle Kilometers Travelled in the Paris Region," *Case Studies on Transport Policy*, Vol. 8, No. 3, pp. 1010–1018, DOI: 10.1016/j.cstp.2020.04.005.

Borenstein, Michael, Larry V. Hedges, Julian P.T. Higgins, and Hannah R. Rothstein (2021) *Introduction to Meta-Analysis*: Wiley, DOI: 10.1002/9781119558378.

Brazil, Noli and David Kirk (2020) "Ridehailing and Alcohol-Involved Traffic Fatalities in the United States: The Average and Heterogeneous Association of Uber," *PLOS ONE*, Vol. 15, No. 9, p. e0238744, DOI: 10.1371/journal.pone.0238744.

Brazil, Noli and David S. Kirk (2016) "Uber and Metropolitan Traffic Fatalities in the United States," *American Journal of Epidemiology*, Vol. 184, No. 3, pp. 192–198, DOI: 10.1093/aje/kww062.

Chen, Yefu (2018) "The Impact of Peer-To-Peer Ridesharing on Travel Mode: Empirical Study of Uber Effects on Travel Mode in Seattle," masters thesis, University of Washington.

Choi, Yunkyung, Subhrajit Guhathakurta, and Anurag Pande (2022) "An Empirical Bayes Approach to Quantifying the Impact of Transportation Network Companies (tncs) Operations on Travel Demand," *Transportation Research Part A: Policy and Practice*, Vol. 161, pp. 269–283, DOI: 10.1016/j.tra.2022.04.008.

Clewlow, Regina R. and Gouri Shankar Mishra (2017) "Disruptive Transportation: The Adoption, Utilization, and Impacts of Ride-Hailing in the United States," Research Report UCD-ITS-RR-17-07, University of California, Davis.

Dean, Brian (2021) "Uber Statistics 2021: How Many People Ride with Uber?," URL: https://web.archive.org/web/20211013213709/https://backlinko.com/uber-users#uber-trips.

Dhanorkar, Suvrat and Gordon Burtch (2022) "The Heterogeneous Effects of P2P Ride-Hailing on Traffic: Evidence from Uber's Entry in California," *Transportation Science*, Vol. 56, No. 3, pp. 750–774, DOI: 10.1287/trsc.2021.1077.

Diao, Mi, Hui Kong, and Jinhua Zhao (2021) "Impacts of transportation network companies on urban mobility," *Nature Sustainability*, DOI: 10.1038/s41893-020-00678-z.

Dills, Angela K. and Sean E. Mulholland (2018) "Ride-Sharing, Fatal Crashes, and Crime," *Southern Economic Journal*, Vol. 84, No. 4, pp. 965–991, DOI: 10.1002/soej.12255.

Erhardt, Gregory D., Jawad Mahmud Hoque, Vedant Goyal, Simon Berrebi, Candace Brakewood, and Kari E. Watkins (2022) "Why Has Public Transit Ridership Declined in the United States?" *Transportation Research Part A: Policy and Practice*, Vol. 161, pp. 68–87, DOI: 10.1016/j.tra.2022.04.006.

Erhardt, Gregory D., Richard Alexander Mucci, Drew Cooper, Bhargava Sana, Mei Chen, and Joe Castiglione (2021) "Do Transportation Network Companies Increase or Decrease Transit Ridership? Empirical Evidence from San Francisco," *Transportation*, DOI: 10.1007/s11116-021-10178-4.

Erhardt, Gregory D., Sneha Roy, Drew Cooper, Bhargava Sana, Mei Chen, and Joe Castiglione (2019) "Do Transportation Network Companies Decrease or Increase Congestion?" *Science Advances*, Vol. 5, No. 5, DOI: 10.1126/sciadv.aau2670.

Fageda, Xavier (2021) "Measuring the Impact of Ride-Hailing Firms on Urban Congestion: The Case of Uber in Europe," *Papers in Regional Science*, Vol. 100, No. 5, pp. 1230–1254, DOI: 10.1111/pirs.12607.

Firpo, Sergio and Vitor Possebom (2018) "Synthetic Control Method: Inference, Sensitivity Analysis and Confidence Sets," *Journal of Causal Inference*, Vol. 6, No. 2, DOI: 10.1515/jci-2016-0026.

Flor, María, Armando Ortuño, and Begoña Guirao (2022a) "Ride-Hailing Services: Competition or Complement to Public Transport to Reduce Accident Rates. The Case of Madrid," *Frontiers in Psychology*, Vol. 13, DOI: 10.3389/fpsyg.2022.951258.

——— (2022b) "Does the Implementation of Ride-Hailing Services Affect Urban Road Safety? The Experience of Madrid," *International Journal of Environmental Research and Public Health*, Vol. 19, No. 5, p. 3078, DOI: 10.3390/ijerph19053078.

Gelman, Andrew and Eric Loken (2013) "The Garden of Forking Paths: Why Multiple Comparisons Can Be a Problem, Even When There Is No "Fishing Expedition" or "p-Hacking" and the Research Hypothesis Was Posited Ahead of Time,"Technical report.

Goodman-Bacon, Andrew (2021) "Difference-in-Differences with Variation in Treatment Timing," *Journal of Econometrics*, Vol. 225, No. 2, pp. 254–277, DOI: 10.1016/jeconom.2021.03.014.

Gorback, Caitlin (2020) "Ridesharing and the Redistribution of Economic Activity," *Working Paper*.

Hall, Jonathan D. (2021) "Can Tolling Help Everyone? Estimating the Aggregate and Distributional Consequences of Congestion Pricing," *Journal of the European Economic Association*, Vol. 19, No. 1, pp. 441–474, DOI: 10.1093/jeea/jvz082.

Hall, Jonathan D., Craig Palsson, and Joseph Price (2018) "Is Uber a Substitute or Complement for Public Transit?" *Journal of Urban Economics*, Vol. 108, No. 1, pp. 36–50, DOI: 10.1016/j.jue.2018.09.003.

Havránek, Tomáš, T. D. Stanley, Hristos Doucouliagos, Pedro Bom, Jerome Geyer-Klingeberg, Ichiro Iwasaki, W. Robert Reed, Katja Rost, and R. C. M. van Aert (2020) "Reporting Guidelines for Meta-Analysis in Economics," *Journal of Economic Surveys*, Vol. 34, No. 3, pp. 469–475, DOI: 10.1111/joes.12363.

Heblich, Stephan, Stephen J Redding, and Daniel M Sturm (2020) "The Making of the Modern Metropolis: Evidence from London," *Quarterly Journal of Economics*, Vol. 135, No. 4, pp. 2059–2133, DOI: 10.1093/qje/qjaa014.

Heme, Mostari Jahan, Atiya Anika, and Sk. Md. Mashrur (2020) "Impact of Ride-Sharing on Public Transport in Dhaka City: An Exploratory Study," in *Proceedings of the 5th International Conference on Civil Engineering for Sustainable Development*, URL: http://iccesd.com/proc_2020/Papers/TRE-4491.pdf.

Henao, Alejandro and Wesley E. Marshall (2018) "The Impact of Ride-Hailing on Vehicle Miles Traveled," *Transportation*, Vol. 46, No. 6, pp. 2173–2194, DOI: 10.1007/s11116-018-9923-2.

Herzog, Ian (2021) "The City-wide Effects of Tolling Downtown Drivers: Evidence from London's Congestion Charge," *Working Paper*.

Huang, Jonathan Yinhao, Farhan Majid, and Mark Daku (2019) "Estimating Effects of Uber Ride-Sharing Service on Road Traffic-Related Deaths in South Africa: A Quasi-Experimental Study," *Journal of Epidemiology and Community Health*, Vol. 73, No. 3, pp. 263–271, DOI: 10.1136/jech-2018-211006.

Kirk, David S., Nicolo Cavalli, and Noli Brazil (2020) "The Implications of Ridehailing for Risky Driving and Road Accident Injuries and Fatalities," *Social Science & Medicine*, Vol. 250, p. 112793, DOI: 10.1016/j.socscimed.2020.112793.

Krishnamurthy, Chandra Kiran B. and Nicole Ngo (2022) "Do Transportation Network Companies Worsen Congestion and Air Quality? Evidence from Uber in California," *Working Paper*.

Leard, Benjamin and Jiawei Xing (2020) "What Does Ridesharing Replace?" Working Paper 20-03, Resources for the Future.

Lee, Kyunghee, Qianran (Jenny) Jin, Animesh Animesh, and Jui Ramaprasad (2022) "Impact of Ride-Hailing Services on Transportation Mode Choices: Evidence

from Traffic and Transit Ridership," *MIS Quarterly*, Vol. 46, No. 4, pp. 1875–1900, DOI: 10.25300/misq/2022/15707.

LeRoy, Stephen F. and Jon Sonstelie (1983) "Paradise Lost and Regained: Transportation Innovation, Income, and Residential Location," *Journal of Urban Economics*, Vol. 13, No. 1, pp. 67–89, DOI: 10.1016/0094-1190(83)90046-3.

Li, Wenting, Amer Shalaby, and Khandker Nurul Habib (2021) "Exploring the Correlation between Ride-Hailing and Multimodal Transit Ridership in Toronto," *Transportation*, Vol. 49, No. 3, pp. 765–789, DOI: 10.1007/s11116-021-10193-5.

Li, Ziru, Chen Liang, Yili Hong, and Zhongju Zhang (2022) "How Do On-demand Ridesharing Services Affect Traffic Congestion? The Moderating Role of Urban Compactness," *Production and Operations Management*, Vol. 31, No. 1, pp. 239–258, DOI: 10.1111/poms.13530.

Malalgoda, Narendra and Siew Hoon Lim (2019) "Do Transportation Network Companies Reduce Public Transit Use in the U.S.?" *Transportation Research Part A: Policy and Practice*, Vol. 130, pp. 351–372, DOI: 10.1016/j.tra.2019.09.051.

Nazif-Munoz, José Ignacio, Brice Batomen, and Arijit Nandi (2022) "Does Ridesharing Affect Road Safety? The Introduction of Moto-Uber and Other Factors in the Dominican Republic," *Research in Globalization*, Vol. 4, p. 100077, DOI: 10.1016/j.resglo.2021.100077.

Nelson, Erik and Nicole Sadowsky (2018) "Estimating the Impact of Ride-Hailing App Company Entry on Public Transportation Use in Major US Urban Areas," *The B.E. Journal of Economic Analysis & Policy*, Vol. 19, No. 1, DOI: 10.1515/bejeap-2018-0151.

Ngo, Nicole S., Thomas Götschi, and Benjamin Y. Clark (2021) "The Effects of Ride-Hailing Services on Bus Ridership in a Medium-Sized Urban Area Using Micro-Level Data: Evidence from the Lane Transit District," *Transport Policy*, Vol. 105, pp. 44–53, DOI: 10.1016/j.tranpol.2021.02.012.

Pan, Yang and Liangfei Qiu (2022) "How Ride-Sharing Is Shaping Public Transit System: A Counterfactual Estimator Approach," *Production and Operations Management*, Vol. 31, No. 3, pp. 906–927, DOI: 10.1111/poms.13582.

Pang, Jindong and Shulin Shen (2022) "Do Ridesharing Services Cause Traffic Congestion?" *Canadian Journal of Economics/Revue canadienne d'économique*, Vol. 56, No. 2, pp. 520–552, DOI: 10.1111/caje.12630.

Qian, Xinwu, Tian Lei, Jiawei Xue, Zengxiang Lei, and Satish V. Ukkusuri (2020) "Impact of Transportation Network Companies on Urban Congestion: Evidence from Large-scale Trajectory Data," *Sustainable Cities and Society*, Vol. 55, DOI: 10.1016/j.scs.2020.102053.

Redman-Ernst, Gilbert Michael (2021) "Effects of Uber on Traffic Fatalities in the United States," masters thesis, Miami University.

Reynolds, Benjamin (2020) "Disruptor of Transportation:Uber's Effect on Public Transportation In England," masters thesis, University of Essex.

Roy, Sneha, Drew Cooper, Alex Mucci, Bhargava Sana, Mei Chen, Joe Castiglione, and Gregory D. Erhardt (2020) "Why Is Traffic Congestion Getting Worse? A Decomposition of the Contributors to Growing Congestion in San Francisco-Determining the Role of Tncs," *Case Studies on Transport Policy*, Vol. 8, No. 4, pp. 1371–1382, DOI: 10.1016/j.cstp.2020.09.008.

Sander, Ryan M. (2019) "Estimating the Causal Effects of Ride-Hailing Technologies on Traffic Congestion," working paper.

Schaller, Bruce (2021) "Can Sharing a Ride Make for Less Traffic? Evidence from Uber and Lyft and Implications for Cities," *Transport Policy*, Vol. 102, pp. 1–10, DOI: 10.1016/j.tranpol.2020.12.015.

Scholl, Lynn, Felipe Bedoya, Orlando Sabogal-Cardona, and Daniel Oviedo (2022) "Making the Links between Ride-Hailing and Public Transit Ridership: Impacts in Medium and Large Colombian Cities," *Research in Transportation Business & Management*, Vol. 45, p. 100901, DOI: 10.1016/j.rtbm.2022.100901.

Shi, Kunbo, Rui Shao, Jonas De Vos, Long Cheng, and Frank Witlox (2021) "The Influence of Ride-Hailing on Travel Frequency and Mode Choice," *Transportation Research Part D: Transport and Environment*, Vol. 101, p. 103125, DOI: 10.1016/j.trd.2021.103125.

Shi, Xiaoyang, Zhengquan Li, and Enjun Xia (2021) "The Impact of Ride-Hailing and Shared Bikes on Public Transit: Moderating Effect of the Legitimacy," *Research in Transportation Economics*, Vol. 85, p. 100870, DOI: 10.1016/j.retrec.2020.100870.

Sturgeon, Lianne Renee (2019) "The Impact of Transportation Network Companies on Public Transit: A Case Study at the San Francisco International Airport," Scripps Senior Theses 1318, Scripps College.

Tarduno, Matthew (2021) "The Congestion Costs of Uber and Lyft," *Journal of Urban Economics*, Vol. 122, p. 103318, DOI: https://doi.org/10.1016/j.jue.2020.103318.

Taylor, Brian D. and Camille Fink (2013) "Explaining Transit Ridership: What Has the Evidence Shown?" *Transportation Letters*, Vol. 5, No. 1, pp. 15–26, DOI: 10.1179/1942786712z.0000000003.

Teltser, Keith, Conor Lennon, and Jacob Burgdorf (2021) "Do Ridesharing Services Increase Alcohol Consumption?" *Journal of Health Economics*, Vol. 77, p. 102451, DOI: 10.1016/j.jhealeco.2021.102451.

Tirachini, Alejandro, Emmanouil Chaniotakis, Mohamed Abouelela, and Constantinos Antoniou (2020) "The Sustainability of Shared Mobility: Can a Platform for Shared Rides Reduce Motorized Traffic in Cities?" *Transportation Research Part C: Emerging Technologies*, Vol. 117, p. 102707, DOI: 10.1016/j.trc.2020.102707.

Tirachini, Alejandro and Mariana del Río (2019) "Ride-hailing in Santiago de Chile: Users' characterisation and effects on travel behaviour," *Transport Policy*, Vol. 82, pp. 46–57, DOI: 10.1016/j.tranpol.2019.07.008.

Uber Technologies, Inc. (2019) "US Safety Report,"Technical report, Uber Technologies, Inc.

Wang, Yiyuan, Anne Vernez Moudon, and Qing Shen (2021) "How Does Ride-Hailing Influence Individual Mode Choice? An Examination Using Longitudinal Trip Data from the Seattle Region," *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2676, No. 3, pp. 621–633, DOI: 10.1177/03611981211055669.

Ward, Jacob W., Jeremy J. Michalek, Inês L. Azevedo, Constantine Samaras, and 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*, Vol. 108, pp. 289–301, DOI: 10.1016/j.trc.2019.07.026.

Ward, Jacob W., Jeremy J. Michalek, Constantine Samaras, Inês L. Azevedo, Alejandro Henao, Clement Rames, and Tom Wenzel (2021) "The Impact of Uber and Lyft on Vehicle Ownership, Fuel Economy, and Transit across U.S. Cities," *iScience*, Vol. 24, No. 1, p. 101933, DOI: 10.1016/j.isci.2020.101933.

Wu, Xiatian and Don MacKenzie (2021) "Assessing the VMT Effect of Ridesourcing Services in the US," *Transportation Research Part D: Transport and Environment*, Vol. 94, p. 102816, DOI: 10.1016/j.trd.2021.102816.

Young, Mischa and Steven Farber (2019) "The Who, Why, and When of Uber and Other Ride-Hailing Trips: An Examination of a Large Sample Household Travel Survey," *Transportation Research Part A: Policy and Practice*, Vol. 119, pp. 383–392, DOI: 10.1016/j.tra.2018.11.018.

Zheng, Emily (2019) "Can Uber and Lyft Save Public Transit?" Pomona Senior Theses 221, Pomona College.

Figure 1: Outcome variables over time for Vancouver and comparable cities
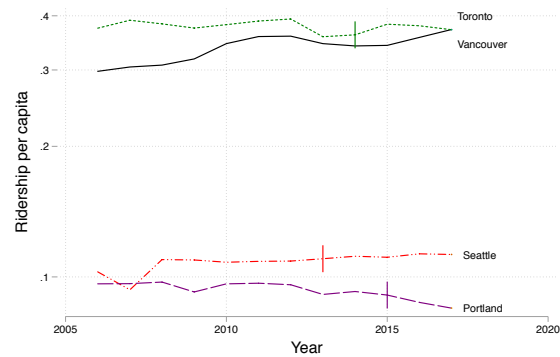
(a) Transit ridership (per capita)
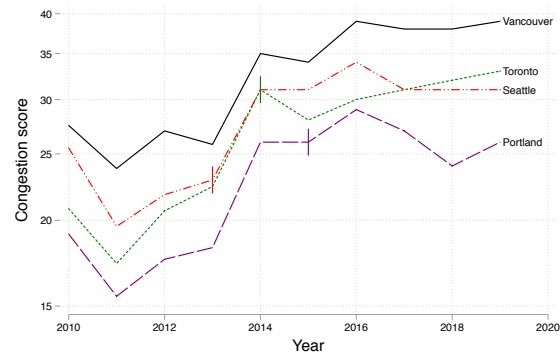
(b) Congestion (combined measure)

(c) Traffic fatalities (per capita)

Notes: Each panel shows outcome variables for Vancouver, BC, Toronto, ON, Seattle, WA, and Portland, OR. Panel (a) shows ridership, (b) shows traffic congestion, and (c) shows traffic fatalities. Uber entered Portland for a few weeks in 2014, but neither Uber or Lyft established a permanent presence until 2015.

Figure 2: Frequency histogram of pairwise difference-in-difference estimates
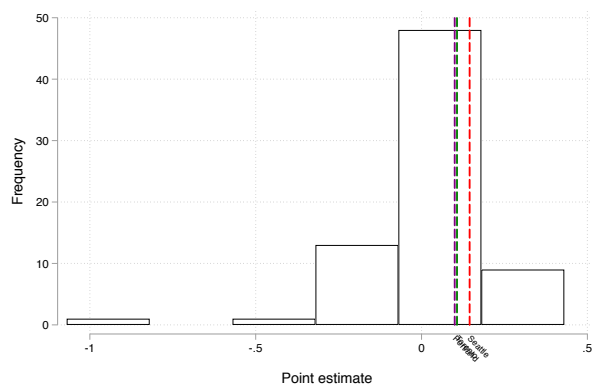
(a) Transit ridership (per capita)
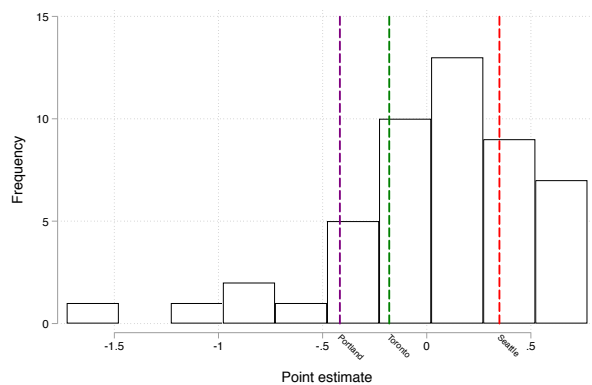
(b) Congestion (combined measure)
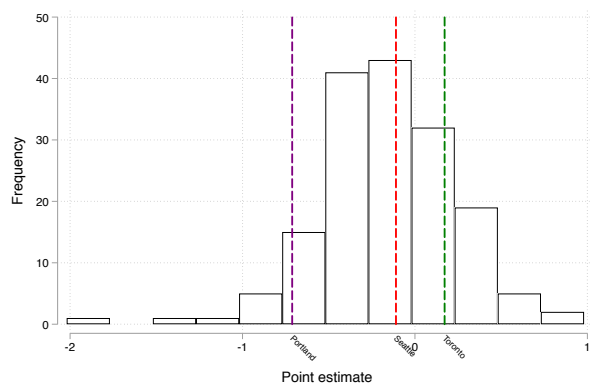
(c) Traffic fatalities (per capita)

Notes: Each panel displays a frequency histogram of pairwise difference-in-difference estimates in a regression that includes Vancouver and one other city in which ride-hailing entered between 2013 and 2015. Panel (a) shows ridership, (b) shows traffic congestion, and (c) shows traffic fatalities. Estimates from regressions including Portland, OR, Seattle, and Toronto are marked with dashed lines.

Figure 3: Estimated treatment effects
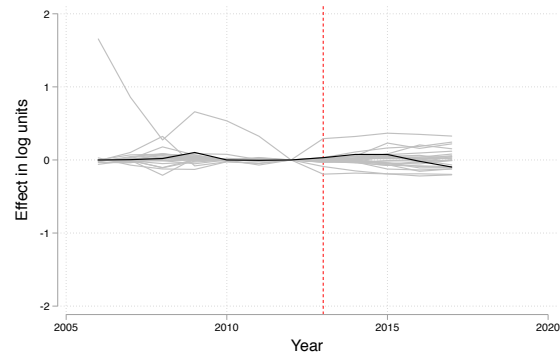
(a) Ridership
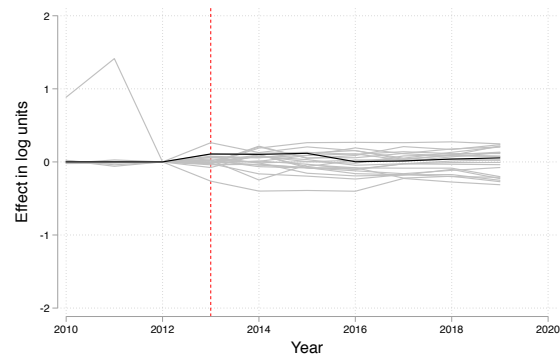
(b) Traffic congestion

(c) Traffic fatalities

Notes: Each panel shows treatment effects for Vancouver (black line) and placebo treatment effects for donors (light gray). Panel (a) shows ridership, (b) shows traffic congestion, and (c) shows traffic fatalities. The donor group is cities where a ride-hailing company began offering services in 2013. For the list of donor cities and their weights, see Appendix Tables A2-A4.

Table 1: Effect of ride-hailing on ridership, congestion, and fatalities

Year	Ridership (1)	Congestion (2)	Fatalities (3)
2013	0.031	0.109	−0.150
	(0.292)	(0.333)	(0.500)
2014	0.076	0.103	0.212
	(0.208)	(0.524)	(0.625)
2015	0.075	0.118	0.583
	(0.500)	(0.524)	(0.167)
2016	−0.019	0.002	0.521
	(0.792)	(0.714)	(0.042)
2017	−0.097	0.013	0.654
	(0.625)	(0.714)	(0.042)
2018		0.039	
		(0.667)	
2019		0.055	
		(0.667)	
Average	0.013	0.063	0.365
(RMSPE std. p)	(0.667)	(0.667)	(0.167)
Overall C.I.	[-0.013,0.040]	[-0.063,0.188]	[-0.365,1.094]

Notes: Standardized p-values are in parentheses and confidence intervals are in brackets. Confidence levels are 96% for ridership, 95% for congestion, and 96% for fatalities. See Footnote 19 for an explanation of how confidence levels are expressed following Firpo and Possebom (2018).

Table 2: Robustness checks

	2013 synth. (1)	Alt. matching (2)	Recess. + alt. match. (3)	Leave-one-out (4)	2014 synth. (5)
Transit ridership	0.013	0.059	0.033	0.008	−0.059
	[−0.013, 0.040]	[−0.059, 0.178]	[−0.033, 0.099]	[−0.028, 0.044]	[−0.176, 0.059]
Congestion	0.063	0.063	—	0.062	0.150
	[−0.063, 0.188]	[−0.063, 0.188]	—	[−0.062, 0.186]	[−0.150, 0.451]
Traffic fatalities	0.365	0.367	0.435	0.416	0.509
	[−0.365, 1.094]	[−0.367, 1.101]	[−0.435, 1.304]	[−0.416, 1.247]	[−0.509, 1.528]

Notes: The first column reproduces our baseline results for comparison. Confidence levels for the first four columns are 96% for ridership, 95% for congestion, and 96% for fatalities. For 2014 "treatment," (Column 5), confidence levels are 94%, 96%, and 95%, respectively. See Footnote 19 for an explanation of how confidence levels are expressed following Firpo and Possebom (2018). Column (4) reports the averages of point estimates and confidence interval bounds over all leave-one-out synthetic controls.

Table 3: Summary of the articles estimating the effect of ride-hailing on transit ridership

Article	Treatment effect & sources of heterogeneity[*]	Setting	Method	Published[†]
Agarwal et al. (2023)	+2.4% subway ridership during ride-hailing strike	India, 2017–2018	Difference-in-differences, regression discontinuity design	Yes
Babar and Burtch (2020)	-1.3% for bus, +3.0% for commuter rail, no effect on subway or light rail. Important sources of heterogeneity are local population, rates of violent crime, weather, gas prices, transit riders' average trip distance, & the overall quality of public transit options	US, 2012–2018	Difference-in-differences	Yes
Chen (2018)	+2.0% on average, +8.8% after five years	Seattle, 2010–2016	Difference-in-differences	No
Clewlow and Mishra (2017)	-6% for bus, -3% for light rail, +3% for commuter rail	US, 2014–2016	Survey analysis	No
Diao et al. (2021)	-8.5% on average, -15% after four years	US, 2005–2016	Panel data regression	Yes
Erhardt et al. (2021)	-10% for bus after five years, No effect on light rail	San Francisco, 2010–2015	Panel data and time-series regressions	Yes
Erhardt et al. (2022)	-10.6% for bus, No effect for rail for large large metropolitan areas, -10% for rail in mid-sized metropolitan areas	US, 2012–2018	Panel data regression	Yes
Hall et al. (2018)	No effect on average, +5% after two years, Larger effects in larger cities and for smaller transit agencies	US, 2004–2015	Difference-in-differences	Yes

Notes: [*]The estimated effect of ride-hailing on transit ridership. Estimates reported in logs were converted to percentage changes. Sources of heterogeneity are reported if the article discusses them in the title, abstract, introduction, or conclusion.
[†]Whether the article is published in a peer-reviewed journal.

Table 3 continued

Article	Treatment effect & sources of heterogeneity[*]	Setting	Method	Published[†]
Heme et al. (2020)	-39% overall	Dhaka City, Bangladesh, 2019	Survey analysis	No
Lee et al. (2022)	No effect on average, but treatment effect is more negative for more compact cities (–5 percentage point difference between effect in New York City & San Diego)	US, 2012–2015	Difference-in-differences	Yes
Li et al. (2022)	No effect on average, +8% for the least compact urbanized areas, -17% for the most compact urbanized areas	US, 2012–2018	Difference-in-differences	Yes
Malalgoda and Lim (2019)	No effect for bus, No average effect for rail, +7% by 2015 for rail	US, 2007–2017	Panel data regression	Yes
Nelson and Sadowsky (2018)	Increases with one ride-hailing firm, No effect or decreases with two	US, 2008–2016	Dynamic-entry event design	Yes
Ngo et al. (2021)	-5.3% for bus	Lane County, Oregon, 2012–2017	Difference-in-differences	Yes
Pan and Qiu (2022)	-7.7% for bus on average, -9.5% for bus after three years	US, 2006–2017	Difference-in-differences and counterfactual estimators	Yes
Reynolds (2020)	-5% overall, -6% for bus, +8% for rail	England, 2004–2018	Panel data regression	No
Scholl et al. (2022)	No significant effect	Columbia, 2005–2018	Difference-in-differences	Yes

Notes: [*]The estimated effect of ride-hailing on transit ridership. Estimates reported in logs were converted to percentage changes. Sources of heterogeneity are reported if the article discusses them in the title, abstract, introduction, or conclusion.
[†]Whether the article is published in a peer-reviewed journal.

Table 3 continued

Article	Treatment effect & sources of heterogeneity[*]	Setting	Method	Published[†]
Shi et al. (2021)	-11% for bus, +146% for rail, Legitimacy increases magnitude of effect	China, 2004–2017	Difference-in-differences	Yes
Shi et al. (2021)	62% of survey respondents say they would use transit more if ride-hailing was not available, around 5% say they would use transit less	Chengdu, China, 2019	Survey analysis	Yes
Sturgeon (2019)	Ridership at airport train station falls by 16 trips per year	San Francisco, 2015–2018	Panel data regression	No
Tirachini and del Río (2019)	"For every [ride-hail] rider that combines with public transport, 11 riders substitute it."	Santiago de Chile, 2017	Survey analysis	Yes
Wang et al. (2021)	No significant effect	Seattle, 2012–2018	Panel data regression	Yes
Ward et al. (2021)	No effect on average, -7% for cities with above median income & above median share of households without children	US, 2011–2017	Difference-in-differences	Yes
Wu and MacKenzie (2021)	Increases transit ridership	US, 2017	Propensity score matching	Yes
Young and Farber (2019)	No significant effect	Toronto, 2016	Survey analysis	Yes
Zheng (2019)	No effect on average, but increases ridership for white collar workers & workers with long hours	US, 2005–2017	Panel data regression	No

Notes: [*]The estimated effect of ride-hailing on transit ridership. Estimates reported in logs were converted to percentage changes. Sources of heterogeneity are reported if the article discusses them in the title, abstract, introduction, or conclusion.
[†]Whether the article is published in a peer-reviewed journal.

Table 4: Summary of the articles estimating the effect of ride-hailing on the Planning Time Index

Article	Treatment effect & sources of heterogeneity[*]	Setting	Method	Published[†]
Lee et al. (2022)	No effect on average, Treatment effect is more positive for more compact cities (9 percentage points larger change in New York than in San Diego)	US, 2012–2015	Difference-in-differences	Yes
Li et al. (2022)	No effect on average, -9.96% (not statistically significant) for the least compact urbanized areas, +20.2% for the most compact urbanized areas	US, 2012–2018	Difference-in-differences	Yes

Notes: [*]The estimated effect of ride-hailing on the Planning Time Index, which is the ratio of the 95th percentile of travel times to the free-flow travel time. Sources of heterogeneity are reported if the article discusses them in the title, abstract, introduction, or conclusion.
[†]Whether the article is published in a peer-reviewed journal.

Table 5: Summary of the articles estimating the effect of ride-hailing on travel speed

Article	Treatment effect & sources of heterogeneity[*]	Setting	Method	Published[†]
Alexander and González (2015)	-1.16% to 7.25%	Boston, 2010	Counterfactual simulation[§]	No
Erhardt et al. (2019)	+9%	San Franscisco, 2010–2016	Panel data regression, counterfactual simulation[§]	Yes
Qian et al. (2020)	-22.5% on weekdays	New York City, 2017–2019	Time series	Yes
Roy et al. (2020)	-3.3%	San Franscisco, 2010–2016	Counterfactual simulation[§]	Yes
Tarduno (2021)	-2.3%	Austin, Texas, 2015–2016	Difference-in-differences	Yes
Krishnamurthy and Ngo (2022)	+3.45% on average, Worsens congestion for the most populated counties & congested times	California, 2009–2017	Difference-in-differences	No

Notes: [*]The estimated effect of ride-hailing on travel speed. Estimates reported in logs were converted to percentage changes. The estimates from Alexander and González (2015) were converted using the fact that vehicle miles traveled divided by vehicle hours traveled is average speed. Sources of heterogeneity are reported if the article discusses them in the title, abstract, introduction, or conclusion.
[†]Whether the article is published in a peer-reviewed journal.
[§]We use the term counterfactual simulation for a variety of methods which use a more formal model, such as a travel demand model, to estimate the counterfactual.

Table 6: Summary of the articles estimating the effect of ride-hailing on the travel time

Article	Treatment effect & sources of heterogeneity[*]	Setting	Method	Published[†]
Hall et al. (2018)	0.6% to +1.3% in MSAs with high population or low transit ridership	US, 2004–2015	Difference-in-differences	Yes
Sander (2019)	No significant effect	Texas, 2019	Panel data regression	No
Basak et al. (2020)	+1.73%	US, 2013–2015	Mediation analysis	No
Diao et al. (2021)	+0.89%	US, 2005–2016	Panel data regression	Yes
Fageda (2021)	-3.5% on average, -7.7% with limited regulation, -12% after 3 years	Europe, 2008–2016	Panel data regression	Yes
Agarwal et al. (2023)	+3.2% to +4.6% on average, As large as +10.1% to +14.8% at peak times	India, 2017–2018	Difference-in-differences, regression discontinuity design	Yes

Notes: [*]The estimated effect of ride-hailing on travel times. Estimates reported in logs were converted to percentage changes.
[†]Whether the article is published in a peer-reviewed journal.

Table 7: Summary of the articles estimating the effect of ride-hailing on the vehicle kilometers traveled

Article	Treatment effect & sources of heterogeneity[*]	Setting	Method	Published[†]
Alexander and González (2015)	-11.57% to +1.83%	Boston, 2010	Counterfactual simulation[§]	No
Henao and Marshall (2018)	+83.5%	Denver, 2016	Quasi-natural experiment	Yes
Erhardt et al. (2019)	+6%	San Franscisco, 2010–2016	Panel data regression, counterfactual simulation[§]	Yes
Ward et al. (2019)	No significant effect on VKT, -3% change in vehicle registrations	US, 2005–2015	Difference-in-differences	Yes
Bekka et al. (2020)	No significant effect	Paris, France, 2018	Survey analysis	Yes
Leard and Xing (2020)	+8% on average, +16% for CBSAs over three million, +2% for CBSAs below one million	US, 2016–2017	Survey analysis	No
Roy et al. (2020)	+7.1%	San Franscisco, 2010–2016	Counterfactual simulation[§]	Yes
Tirachini et al. (2020)	Shared rides in cars increase VKT +51% to +96%, Shared rides in vans reduce VKT	Mexico City, 2019	Survey analysis	Yes
Schaller (2021)	Average ride-hailing trip doubles the net counterfactual VKT	US, 2014–2020	Counterfactual model	Yes
Wu and MacKenzie (2021)	+2.95%	US, 2017	Propensity score matching	Yes
Choi et al. (2022)	+5.9% on average, +5.4% after 6 years	Atlanta, 2012–2018	Empirical Bayes model	Yes

Notes: [*]The estimated effect of ride-hailing on vehicle kilometers traveled (VKT). Estimates reported in logs were converted to percentage changes. Sources of heterogeneity are reported if the article discusses them in the title, abstract, introduction, or conclusion.
[†]Whether the article is published in a peer-reviewed journal.
[§]We use the term counterfactual simulation for a variety of methods which use a more formal model, such as a travel demand model, to estimate the counterfactual.

Table 7 continued

Article	Treatment effect & sources of heterogeneity[*]	Setting	Method	Published[†]
Dhanorkar and Burtch (2022)	No effect on average, but with significant heterogeneity. For example, -1.5% on weekdays in areas with low population density, +8.2% on weekends on local roads, +8.5% to +8.8% in areas with higher prior public transport usage	California, 2010–2015	Difference-in-differences & propensity score matching	Yes
Krishnamurthy and Ngo (2022)	+8% on average, +18% for most populated counties	California, 2009–2017	Difference-in-differences	No
Pang and Shen (2022)	No effect on average, -2% on highways, +4% on collector roads	US, 2011–2017	Difference-in-differences	Yes

Notes: [*]The estimated effect of ride-hailing on vehicle kilometers traveled (VKT). Estimates reported in logs were converted to percentage changes. Sources of heterogeneity are reported if the article discusses them in the title, abstract, introduction, or conclusion.
[†]Whether the article is published in a peer-reviewed journal.

Table 8: Summary of the articles estimating the effect of ride-hailing on traffic fatalities

Article	Treatment effect & sources of heterogeneity[*]	Setting	Method	Published[†]
Anderson and Davis (2021)	-4%	US, 2001–2016	Difference-in-differences	Yes
Barreto et al. (2021)	-10% on average, -16% after five quarters	Brazil, 2011–2016	Difference-in-differences	Yes
Barrios et al. (2022)	+3.4% on average, +15% after four years	US, 2001–2016	Difference-in-differences	Yes
Brazil and Kirk (2016)	No significant effect	US, 2015–2014	Difference-in-differences	Yes
Brazil and Kirk (2020)	No effect in aggregate, +6% in most densely populated counties	US, 2009–2017	Difference-in-differences	Yes
Dills and Mulholland (2018)	-1% average change in fatal crashes, -17% to -40% after four or more years	US, 2007–2015	Difference-in-differences	Yes
Flor et al. (2022b)	-22% change in fatalities & serious injuries	Madrid, 2013–2019	Panel data regression	Yes
Flor et al. (2022a)	-25% change in fatalities & serious injuries	Madrid, 2013–2019	Panel data regression	Yes
Huang et al. (2019)	No significant effect	South Africa, 2010–2014	Difference-in-differences	Yes
Kirk et al. (2020)	No significant effect on fatalities, -9% serious injuries	Great Britain, 2009–2017	Negative binomial panel regressions	Yes
Nazif-Munoz et al. (2022)	-1% change in fatalities from UberMOTO	Dominican Republic, 2012–2018	Interrupted time-series	Yes
Redman-Ernst (2021)	No significant effect	US, 2005–2018	Difference-in-differences	No

Notes: [*]The estimated effect of ride-hailing on traffic fatalities. Estimates reported in logs were converted to percentage changes. Sources of heterogeneity are reported if the article discusses them in the title, abstract, introduction, or conclusion.
[†]Whether the article is published in a peer-reviewed journal.

Table 9: Summary statistics for articles in meta-analysis

	All	Ridership	Congestion	Fatalities
Published in a peer-reviewed journal	0.79	0.77	0.79	0.92
Average effect is not statistically significant	0.34	0.35	0.33	0.42
Provides long-run estimates	0.21	0.19	0.13	0.33
Estimates heterogeneous treatment effects	0.29	0.31	0.42	0.08
Geographic area				
Africa	0.02	0	0	0.08
Asia	0.07	0.15	0.04	0
Europe	0.11	0.04	0.08	0.25
South America	0.05	0.08	0	0.08
North America, excluding the US	0.05	0.04	0.04	0.08
United States	0.70	0.69	0.83	0.50
Empirical method				
Difference-in-differences	0.43	0.42	0.38	0.67
Other panel data	0.25	0.31	0.17	0.25
Not panel data	0.32	0.27	0.46	0.08
Observations	56	26	24	12

Notes: There are six papers that report results for both transit ridership and traffic congestion, and so the number of observations in the first column is less than the sum of the observations for the other three columns. For empirical methods, many articles use multiple methods. However, we have encoded the empirical method so that "Not panel data" contains those papers not using any panel data methods.

Table 10: Distribution of average effect sizes

	Min (%)	25th (%)	50th (%)	75th (%)	Max (%)	Mean (%)	N
		Percentile					
Transit ridership							
All	-38.9	-6.0	-0.0	2.1	146.0	2.2	29
Bus	-11.3	-7.7	-5.6	-0.0	3.1	-4.2	10
Rail	-3.0	0.4	2.7	3.8	146.0	16.3	10
Total	-38.9	-8.1	-2.0	0.4	2.1	-6.5	9
Traffic congestion							
Travel time	-4.5	-2.0	0.2	2.5	29.0	2.8	12
Vehicle kilometers traveled	-0.9	-0.5	1.5	7.0	83.5	14.6	12
Traffic fatalities	-25.3	-7.0	-0.2	2.3	3.3	-4.4	12

Notes: This table reports statistics summarizing the distribution of estimates of the treatment effect of ride-hailing on the outcomes listed in the first column.

Table 11: Estimates of average effect from meta-analysis

	Effect size (%)	95% confidence interval	p-value	N
Transit ridership				
All	-1.687	(–3.801, 0.427)	.118	21
Bus	-4.291	(–7.246, –1.337)	.004	8
Rail	2.347	(0.011, 4.683)	.049	7
Total	-2.36	(–6.278, 1.558)	.238	6
Traffic congestion				
Travel time	-.943	(–2.679, 0.794)	.287	9
VKT	-.392	(–0.920, 0.135)	.145	3
Traffic fatalities	-4.123	(–9.649, 1.402)	.144	12

Notes: This table reports the results of a random-effects meta-analysis. The mean effect size is a weighted average of all the estimates, where estimates are weighted by the inverse of the sum of the (estimated) variance of the estimate and the variance between studies.

Table 12: Estimates of sources of heterogeneity in effect sizes from meta-analysis

	All (%)	Ridership (%)	Congestion (%)	Fatalities (%)
Published in peer-reviewed journal	-3.688+	-6.081+	1.592	-0.217
	(2.224)	(3.543)	(1.503)	(8.861)
Estimate of long-run effect	1.645	5.812*	-6.359	-5.284
	(2.596)	(2.933)	(4.495)	(7.306)
Paper estimates heterogeneous treatment effects	1.980	3.965+	1.243	-0.352
	(2.337)	(2.362)	(1.481)	(11.36)
Geographic area				
United States	–	–	–	–
Europe	-6.885*	-3.765	-5.926*	-18.11*
	(2.959)	(4.402)	(2.708)	(7.949)
Other	-4.035	-3.582	5.801	-7.842
	(2.781)	(2.973)	(41.42)	(7.619)
Empirical method				
Difference-in-differences	–	–	–	–
Other panel data	-1.425	-2.250	3.689*	–
	(2.463)	(2.128)	(1.866)	
Not panel data	3.305	–	5.768*	6.101
	(4.108)		(2.476)	(11.99)
Observations	53	24	13	16

Notes: This table reports the results of a random-effects meta-regression. We regress the point estimates on the listed possible sources of heterogeneity. Each estimate is weighted by the inverse of the sum of the (estimated) variance of the estimate and the variance between studies. We include but do not report fixed effects for each outcome type (e.g., "bus ridership," "travel times", "fatalities").

Table 13: Share of articles published in a peer-reviewed journal by whether the average effect is statistically significant

Sample	Share published in a peer-reviewed journal			
	All	Ridership	Congestion	Fatalities
Average effect is not statistically significant	0.84	0.89	0.88	0.80
Average effect is statistically significant	0.76	0.71	0.75	1.00
Difference	0.09	0.18	0.13	-0.20
p-value	0.73	0.38	0.63	0.42
Observations	56	26	24	12

Notes: p-value calculated using Fisher's exact test of independence.

Table 14: Share of articles testing for heterogeneous treatment effects by whether the average effect is statistically significant

Sample	Share testing for heterogeneous effects			
	All	Ridership	Congestion	Fatalities
Average effect is not statistically significant	0.42	0.56	0.63	0.20
Average effect is statistically significant	0.20	0.18	0.31	0.00
Difference	0.22	0.38	0.31	0.20
p-value	0.14	0.08	0.20	0.42
Observations	56	26	24	12

Notes: p-value calculated using Fisher's exact test of independence.

A Additional figures and tables

Figure A1: Reported ridership by APTA and Translink, by mode and total
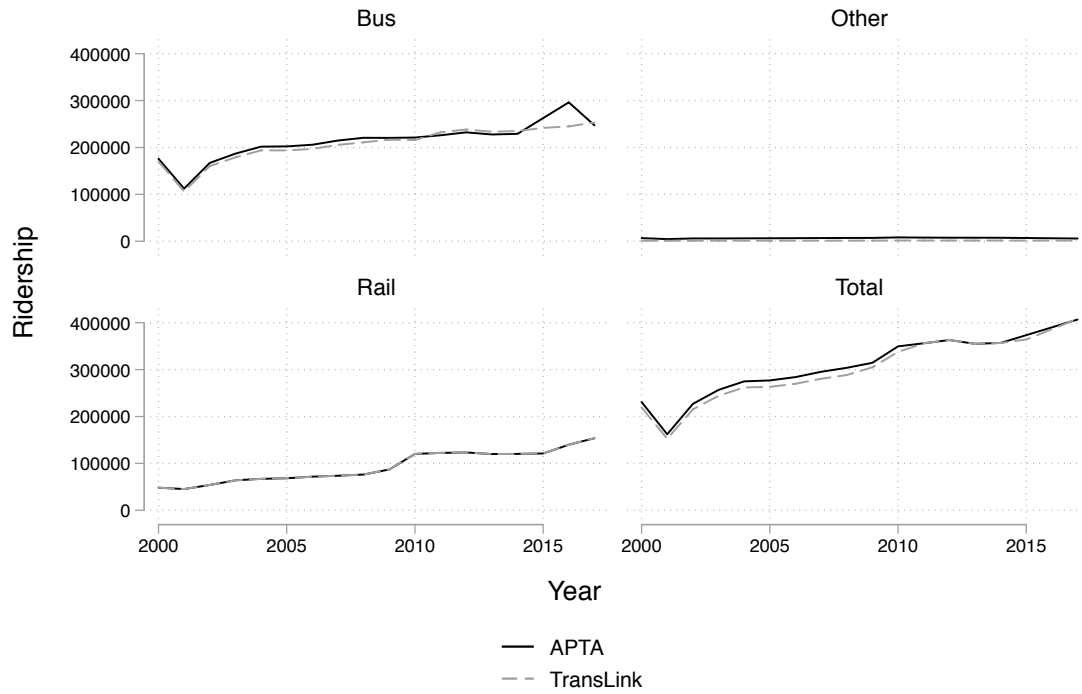
Figure A2: Outcome variables over time for Vancouver and potential donors
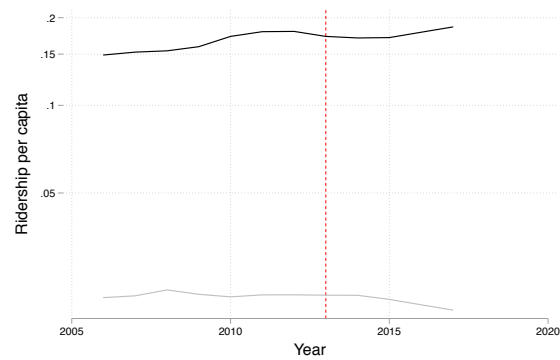
(a) Transit ridership (per capita)
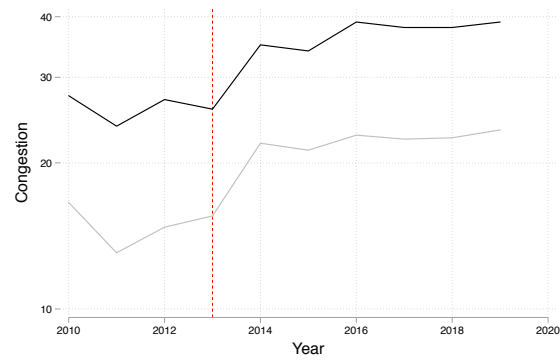
(b) Congestion (combined measure; raw score)
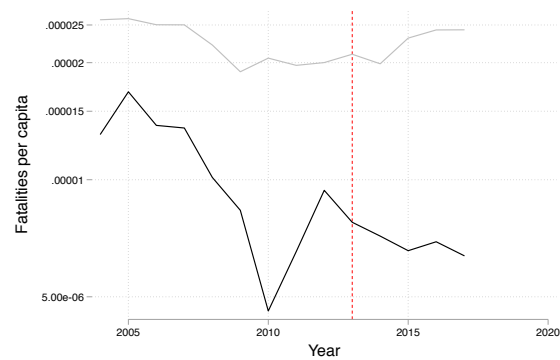
(c) Traffic fatalities (per capita)

Notes: Each panel shows outcome variables for Vancouver (black line) and donors (light gray). Panel (a) shows ridership, (b) shows traffic congestion, and (c) shows traffic fatalities. The potential donor group is cities where a ride-hailing company began offering services in 2013.

Figure A3: Estimated treatment effects with 2014 treatment
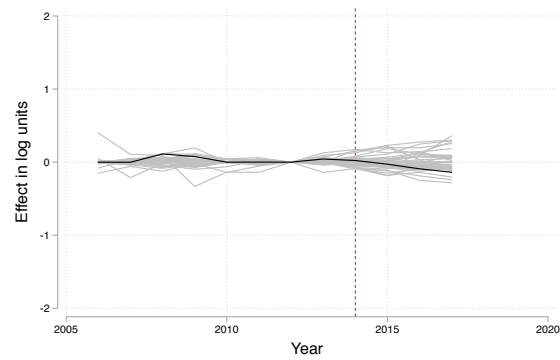
(a) Ridership
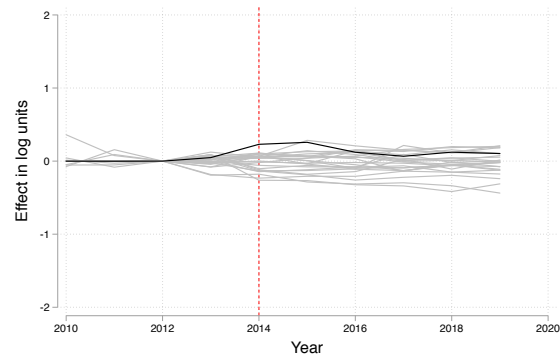
(b) Congestion
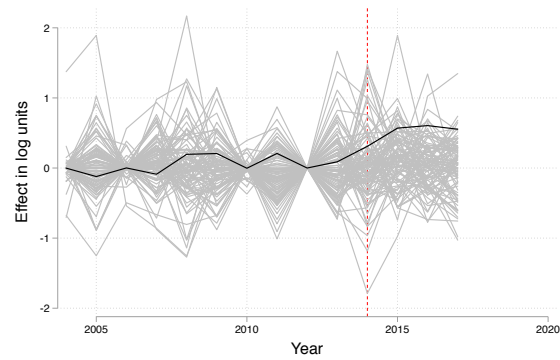
(c) Fatalities

Notes: Each panel shows treatment effects for Vancouver (black line) and placebo treatment effects for donors (light gray). Panel (a) shows ridership, (b) shows traffic congestion, and (c) shows traffic fatalities. The donor group is cities where a ride-hailing company began offering services in 2014.

Figure A4: Ridership treatment effects by mode
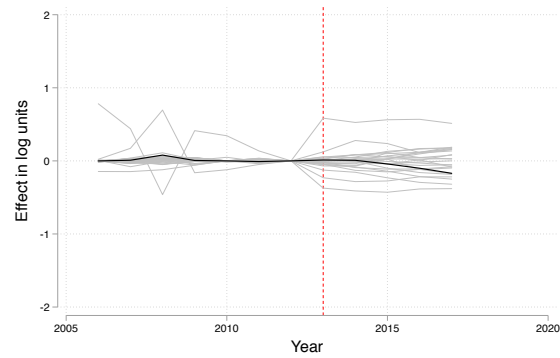
(a) Bus
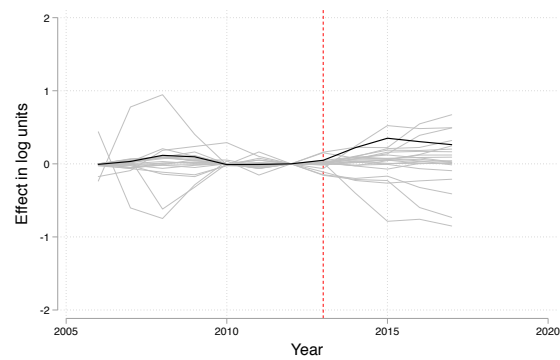
(b) Rail

Notes: Each panel shows ridership for Vancouver (black line) and donors (light gray). Panel (a) shows ridership by bus, and (b) shows rail. The potential donor group is cities with that mode where a ride-hailing company began offering services in 2013.

Figure A5: PRISMA 2020 flow diagram for meta-analysis

57
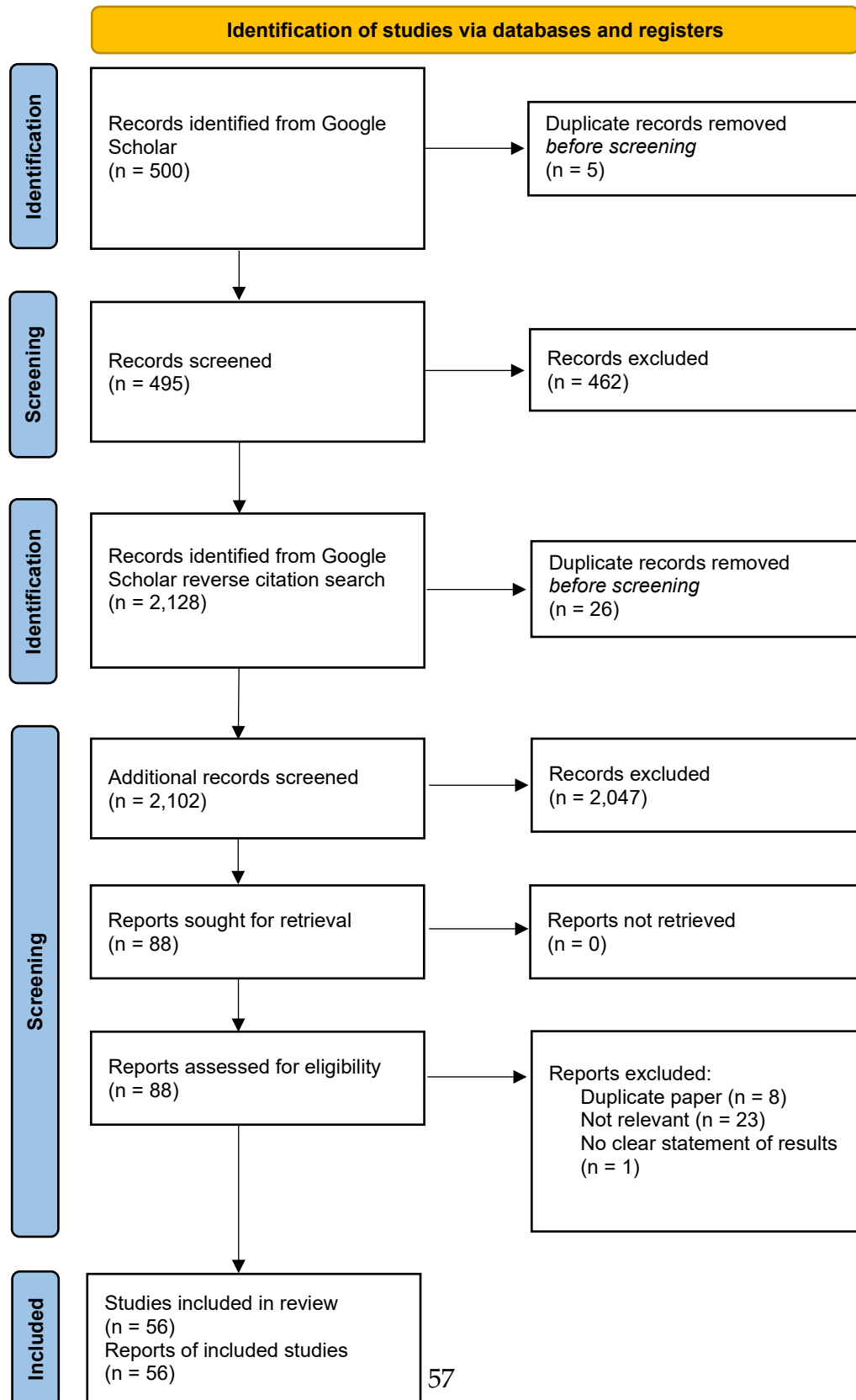
Table A1: Simple difference-in-difference estimates

	Seattle	Vancouver	Toronto	Vancouver	Portland	Vancouver
Panel A: Transit ridership						
Before	-0.2%	3.9%	-0.6%	1.7%	-0.6%	1.4%
After	0.7%	0.7%	0.9%	1.9%	-2.9%	2.9%
Difference	0.9%	-3.2%	1.6%	-0.2%	-2.3%	1.5%
Difference-in-differences	4.1%		1.8%		-3.8%	
Panel B: Traffic congestion						
Before	-7.5%	-0.9%	2.5%	-2.1%	8.0%	6.2%
After	5.2%	5.4%	6.7%	7.1%	0.0%	2.2%
Difference	12.7%	6.3%	4.2%	9.2%	-8.0%	-4.0%
Difference-in-differences	6.4%		-5.0%		-4.0%	
Panel C: Traffic fatalities						
Before	-4.6%	-4.1%	-7.0%	-5.85%	-7.0%	-5.9%
After	0.4%	-7.5%	-0.1%	-4.8%	29.7%	-3.8%
Difference	5.0%	-3.4%	4.5%	0.8%	36.7%	2.1%
Difference-in-differences	8.4%		3.7%		34.6%	

Notes: This table reports difference-in-difference results comparing average growth rates before and after ride-hailing was available in Seattle, Toronto, and Portland to growth rates in Vancouver over the same time period. Average growth rates are calculated using the geometric mean. For Seattle's transit ridership, we take the mean growth rate starting in 2009, rather than 2005, as its growth rate was remarkably consistent from 2009–2012 and so this seems the reasonable comparison.

Table A2: Ridership results by mode

Year	Bus (1)	Rail (2)
2013	0.012	0.048
	(0.792)	(0.190)
2014	0.007	0.221
	(0.958)	(0.238)
2015	−0.042	0.350
	(0.750)	(0.190)
2016	−0.100	0.303
	(0.583)	(0.286)
2017	−0.170	0.263
	(0.375)	(0.333)
Average	−0.059	0.237
(RMSPE std. p)	(0.583)	(0.286)
Overall C.I.	$[-0.176, 0.059]$	$[-0.237, 0.712]$

Notes: Standardized *p*-values are in parentheses and confidence intervals are in brackets. Confidence levels are 96% for Bus ridership and 95% for Rail ridership. See Footnote 19 for an explanation of how confidence levels are expressed following Firpo and Possebom (2018).

Table A3: Distribution of estimates of the effect of ride-hailing on transit ridership from meta-analysis

		Percentile					
	Min (%)	25th (%)	50th (%)	75th (%)	Max (%)	Mean (%)	N
All							
Average	-38.9	-6.0	-0.0	2.1	146.0	2.2	29
Long-run	-9.5	-1.5	6.0	7.2	8.8	2.2	5
Bus							
Average	-11.3	-7.7	-5.6	-0.0	3.1	-4.2	10
Long-run	-9.5	-9.5	-5.5	-1.5	-1.5	-5.5	2
Methodology							
Difference-in-differences	-11.3	-7.7	-5.3	-1.3	1.5	-4.8	5
Other panel data	-8.6	-8.6	-6.4	-0.0	-0.0	-5.0	3
Not panel data	-6.0	-6.0	-1.4	3.1	3.1	-1.4	2
Geography							
United States	-8.6	-6.8	-3.3	0.8	3.1	-3.0	8
Europe	-6.4	-6.4	-6.4	-6.4	-6.4	-6.4	1
Other	-11.3	-11.3	-11.3	-11.3	-11.3	-11.3	1
Rail							
Average	-3.0	0.4	2.7	3.8	146.0	16.3	10
Long-run	7.2	7.2	7.2	7.2	7.2	7.2	1
Methodology							
Difference-in-differences	0.4	1.4	2.7	74.5	146.0	37.9	4
Other panel data	-1.6	-1.6	3.8	7.9	7.9	3.4	3
Not panel data	-3.0	-3.0	1.3	3.0	3.0	0.4	3
Geography							
United States	-3.0	-1.6	1.3	3.0	3.8	1.0	7
Europe	7.9	7.9	7.9	7.9	7.9	7.9	1
Other	2.4	2.4	74.2	146.0	146.0	74.2	2
Total							
Average	-38.9	-8.1	-2.0	0.4	2.1	-6.5	9
Long-run	6.0	6.0	7.4	8.8	8.8	7.4	2
Methodology							
Difference-in-differences	-8.1	-3.9	0.3	1.2	2.0	-1.3	4
Other panel data	-8.5	-8.5	-5.4	-2.0	-2.0	-5.3	3
Not panel data	-38.9	-38.9	-18.4	2.1	2.1	-18.4	2
Geography							
United States	-8.5	-2.0	0.3	2.0	2.1	-1.0	6
Europe	-5.4	-5.4	-5.4	-5.4	-5.4	-5.4	1
Other	-38.9	-38.9	-23.5	-8.1	-8.1	-23.5	2

Notes: This table reports statistics summarizing the distribution of estimates of the treatment effect of ride-hailing on the outcomes listed in the first column. The subgroups for methodology and geography are limited to estimates of the average treatment effect.

Table A4: Distribution of estimates of the effect of ride-hailing on traffic congestion from meta-analysis

	Min (%)	Percentile 25th (%)	50th (%)	75th (%)	Max (%)	Mean (%)	N
Travel time							
Average	-4.5	-2.0	0.2	2.5	29.0	2.8	12
Long-run	-11.0	-11.0	-3.8	3.4	3.4	-3.8	2
Methodology							
Difference-in-differences	-4.5	-3.3	-0.2	0.3	3.4	-0.9	5
Other panel data	-3.4	-1.7	0.4	5.4	9.9	1.8	4
Not panel data	-0.6	-0.6	1.7	29.0	29.0	10.1	3
Geography							
United States	-4.5	-0.6	0.2	1.7	29.0	3.3	10
Europe	-3.4	-3.4	-3.4	-3.4	-3.4	-3.4	1
Other	3.4	3.4	3.4	3.4	3.4	3.4	1
Vehicle kilometers traveled							
Average	-0.9	-0.5	1.5	7.0	83.5	14.6	12
Long-run	5.4	5.4	6.2	7.1	7.1	6.2	2
Methodology							
Difference-in-differences	-0.8	-0.8	-0.5	3.8	8.0	1.5	4
Other panel data	6.0	6.0	6.0	6.0	6.0	6.0	1
Not panel data	-0.9	0.1	3.0	71.0	83.5	23.2	7
Geography							
United States	-0.9	-0.7	1.5	6.0	83.5	10.4	10
Europe	0.2	0.2	0.2	0.2	0.2	0.2	1
Other	71.0	71.0	71.0	71.0	71.0	71.0	1

Notes: See notes for Table A3.

Table A5: Distribution of estimates of the effect of ride-hailing on traffic fatalities from meta-analysis

| | | Percentile | | | | | |
	Min (%)	25th (%)	50th (%)	75th (%)	Max (%)	Mean (%)	N
Average	-25.3	-7.0	-0.2	2.3	3.3	-4.4	12
Long-run	-39.7	-28.0	-10.6	4.6	14.0	-11.7	4
Methodology							
Difference-in-differences	-9.9	-2.2	0.5	2.3	3.3	-0.7	8
Other panel data	-25.3	-25.3	-24.8	3.0	3.0	-15.7	3
Not panel data	-0.4	-0.4	-0.4	-0.4	-0.4	-0.4	1
Geography							
United States	-4.0	-0.3	1.5	2.6	3.3	0.8	6
Europe	-25.3	-25.3	-24.8	3.0	3.0	-15.7	3
Other	-9.9	-9.9	-0.4	-0.0	-0.0	-3.4	3

Notes: See notes for Table A3.

Table A6: Estimates of average effect size in studies set in the United States

	Effect size (%)	95% confidence interval	p-value	N
Transit ridership				
All	-1.603	(–3.789, 0.582)	.15	14
Bus	-3.369	(–6.717, –0.021)	.049	6
Rail	1.048	(–1.397, 3.494)	.401	4
Total	-1.498	(–6.252, 3.256)	.537	4
Traffic congestion				
Travel time	-.546	(–2.323, 1.231)	.547	7
VKT	-.392	(–0.920, 0.135)	.145	3
Traffic fatalities	.604	(–1.856, 3.063)	.63	6

Notes: This table reports the results of a random-effects meta-analysis, limited to studies set in the United States. The mean effect size is a weighted average of all the estimates, where estimates are weighted by the inverse of the sum of the (estimated) variance of the estimate and the variance between studies.

Table A7: Donor weights for 2013 "treatment": Ridership

Agency	CBSA	Baseline	Alt. matching	Great Recession
Metrop. Atlanta Rapid Transit Auth.	Atlanta-Sandy Springs-Marietta, GA	0.013	0	0
Maryland Transit Admin.	Baltimore-Towson, MD	0.023	0	0
Massachusetts Bay Transport. Auth.	Boston-Cambridge-Quincy, MA-NH	0.021	0	0
Charlotte Area Transit System	Charlotte-Gastonia-Rock Hill, NC-SC	0.026	0	0
Chicago Transit Auth.	Chicago-Joliet-Naperville, IL-IN-WI	0.022	0	0
Metra Rail	Chicago-Joliet-Naperville, IL-IN-WI	0.015	0	0
Northern Indiana Commuter Transport. Dist.	Chicago-Joliet-Naperville, IL-IN-WI	0.013	0	0
Trinity Railway Express	Dallas-Fort Worth-Arlington, TX	0.006	0	0
Denver Regional Transport. Dist.	Denver-Aurora-Broomfield, CO	0.012	0.498	0.187
Detroit Transport. Corp.	Detroit-Warren-Livonia, MI	0.019	0	0.324
Metro	Los Angeles-Long Beach-Santa Ana, CA	0.017	0	0
Metrolink	Los Angeles-Long Beach-Santa Ana, CA	0.017	0	0
Metro Transit	Minneapolis-St. Paul-Bloomington, MN-WI	0.016	0	0
Metrop. Transit Auth.	Nashville-Davidson–Murfreesboro–Franklin, TN	0.015	0	0
Regional Transport. Auth.	Nashville-Davidson–Murfreesboro–Franklin, TN	0.017	0.090	0.010
Central Oklahoma Transport. and Parking Auth.	Oklahoma City, OK	0.204	0.412	0
City of Phoenix (Valley Metro)	Phoenix-Mesa-Glendale, AZ	0.015	0	0
San Diego Metrop. Transit System	San Diego-Carlsbad-San Marcos, CA	0.430	0	0.295
Santa Clara Valley Transport. Auth.	San Jose-Sunnyvale-Santa Clara, CA	0.014	0	0
Santa Barbara Metrop. Transit Dist.	Santa Barbara-Santa Maria-Goleta, CA	0.014	0	0
King County Dept. of Transport.	Seattle-Tacoma-Bellevue, WA	0.015	0	0
City of Tucson	Tucson, AZ	0.024	0	0
Virginia Railway Express	Washington-Arlington-Alexandria, DC-VA-MD-WV	0.021	0	0.183
Metro	Washington-Arlington-Alexandria, DC-VA-MD-WV	0.013	0	0

Table A8: Donor weights for 2013 "treatment": Traffic congestion

CBSA	Baseline	Alt. matching
Atlanta-Sandy Springs-Marietta, GA	0.006	0.006
Baltimore-Towson, MD	0.005	0.005
Boston-Cambridge-Quincy, MA-NH	0.363	0.363
Charlotte-Gastonia-Rock Hill, NC-SC	0.006	0.006
Chicago-Joliet-Naperville, IL-IN-WI	0.005	0.005
Dallas-Fort Worth-Arlington, TX	0.011	0.011
Denver-Aurora-Broomfield, CO	0.540	0.540
Detroit-Warren-Livonia, MI	0.003	0.003
Indianapolis-Carmel, IN	0.005	0.005
Los Angeles-Long Beach-Santa Ana, CA	0.007	0.007
Minneapolis-St. Paul-Bloomington, MN-WI	0.004	0.004
Nashville-Davidson–Murfreesboro–Franklin, TN	0.006	0.006
Oklahoma City, OK	0	0
Phoenix-Mesa-Glendale, AZ	0.003	0.003
Providence-New Bedford-Fall River, RI-MA	0.011	0.011
Sacramento–Arden-Arcade–Roseville, CA	0.003	0.003
San Diego-Carlsbad-San Marcos, CA	0.005	0.005
San Jose-Sunnyvale-Santa Clara, CA	0.006	0.006
Seattle-Tacoma-Bellevue, WA	0.004	0.004
Tucson, AZ	0	0
Washington-Arlington-Alexandria, DC-VA-MD-WV	0.007	0.007

Table A9: Donor weights for 2013 "treatment": Traffic fatalities

City	Baseline	Alt. matching	Great Recession
Atlanta, GA	0	0	0
Baltimore, MD	0	0	0
Boston, MA	0.164	0.157	0
Charlotte, NC	0	0	0.239
Chicago, IL	0	0	0
Dallas, TX	0	0	0
Denver, CO	0	0	0
Detroit, MI	0	0	0
Fort Worth, TX	0	0	0
Indianapolis, IN	0	0	0
Los Angeles, CA	0	0	0
Minneapolis, MN	0	0	0
Nashville, TN	0	0	0
Oklahoma City, OK	0	0	0
Phoenix, AZ	0	0	0
Providence, RI	0	0	0
Sacramento, CA	0	0	0
San Diego, CA	0	0	0
San Jose, CA	0	0	0
Santa Barbara, CA	0	0	0
Seattle, WA	0	0	0
St. Paul, MN	0.836	0.843	0.761
Tucson, AZ	0	0	0
Washington, DC	0	0	0

Table A10: Donor weights for 2014 "treatment": Ridership

Agency	CBSA	Weight
New Mexico Dept. of Transport.	Albuquerque, NM	0.090
Golden Empire Transit Dist.	Bakersfield-Delano, CA	0.007
Stark Area RTA	Canton-Massillon, OH	0.003
Lee County Transit	Cape Coral-Fort Myers, FL	0.002
Chattanooga Area RTA	Chattanooga, TN-GA	0.006
Southwest Ohio RTA	Cincinnati-Middletown, OH-KY-IN	0.003
Greater Cleveland RTA	Cleveland-Elyria-Mentor, OH	0
Mountain Metrop. Transit	Colorado Springs, CO	0.002
Central Ohio Transit Auth.	Columbus, OH	0.009
Chapel Hill Transit	Durham-Chapel Hill, NC	0.212
Mass Transp. Auth.	Flint, MI	0.002
Fresno Area Express	Fresno, CA	0
Greensboro Transit Auth.	Greensboro-High Point, NC	0.008
CTTRANSIT (Hartford)	Hartford-West Hartford-East Hartford, CT	0.007
Harris County MTA, Texas	Houston-Sugar Land-Baytown, TX	0.003
Jacksonville Transp. Auth.	Jacksonville, FL	0.005
Capital Area Transp. Auth.	Lansing-East Lansing, MI	0.008
River City Transit Auth.	Louisville/Jefferson County, KY-IN	0.004
Memphis Area Transit Auth.	Memphis, TN-MS-AR	0.007
Miami-Dade Transit	Miami-Fort Lauderdale-Pompano Beach, FL	0.004
South Florida Regional Transp. Auth.	Miami-Fort Lauderdale-Pompano Beach, FL	0.008
Milwaukee County Transit System	Milwaukee-Waukesha-West Allis, WI	0.004
OC Transpo	Ottawa, ON	0.006
Delaware Transit Corp.	Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	0.011
Southeastern Pennsylvania Transp. Auth.	Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	0.008
Port Auth. of Allegheny County	Pittsburgh, PA	0.005
Northern New England Passenger Rail Auth.	Portland-South Portland-Biddeford, ME	0.155
Riverside Transit Agency	Riverside-San Bernardino-Ontario, CA	0.006
Salem-Keizer Transit	Salem, OR	0.003
Utah Transit Auth.	Salt Lake City, UT	0.369
Spokane Transit Auth.	Spokane, WA	0.012
Hillsborough Area Regional Transit Auth.	Tampa-St. Petersburg-Clearwater, FL	0.008
Toronto Transit Commission	Toronto, ON	0.007
Hampton Roads Transit	Virginia Beach-Norfolk-Newport News, VA-NC	0.003
Visalia City Coach	Visalia-Porterville, CA	0.007

Table A11: Donor weights for 2014 "treatment": Traffic congestion

CBSA	Weight
Cincinnati-Middletown, OH-KY-IN	0.001
Cleveland-Elyria-Mentor, OH	0
Columbus, OH	0.001
Hartford-West Hartford-East Hartford, CT	0.562
Honolulu, HI	0
Houston-Sugar Land-Baytown, TX	0.001
Jacksonville, FL	0.001
Kansas City, MO-KS	0.001
Louisville/Jefferson County, KY-IN	0.416
Memphis, TN-MS-AR	0.001
Miami-Fort Lauderdale-Pompano Beach, FL	0.002
Milwaukee-Waukesha-West Allis, WI	0.004
Montreal, QC	0
Orlando-Kissimmee-Sanford, FL	0.001
Ottawa, ON	0.001
Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	0.001
Pittsburgh, PA	0.001
Raleigh-Cary, NC	0.001
Richmond, VA	0.001
Riverside-San Bernardino-Ontario, CA	0.001
Salt Lake City, UT	0
Tampa-St. Petersburg-Clearwater, FL	0.003
Toronto, ON	0.002
Virginia Beach-Norfolk-Newport News, VA-NC	0.001

Table A12: Donor weights for 2014 "treatment": Traffic fatalities

City	Weight	City	Weight
Akron, OH	0.006	Lexington-Fayette, KY	0.005
Albuquerque, NM	0.005	Lincoln, NE	0.007
Amarillo, TX	0.004	Little Rock, AR	0.004
Asheville, NC	0.003	Louisville, KY	0.004
Athens-Clarke, GA	0.002	Lubbock, TX	0.003
Bakersfield, CA	0.004	Madison, WI	0.004
Baton Rouge, LA	0.003	Manchester, NH	0.002
Bloomington, IN	0.002	Memphis, TN	0.004
Boise City, ID	0.003	Miami, FL	0.005
Cape Coral, FL	0.005	Milwaukee, WI	0.003
Cedar Rapids, IA	0.003	Modesto, CA	0.005
Charleston, SC	0.006	Myrtle Beach, SC	0.004
Chattanooga, TN	0.004	New Haven, CT	0.003
Cincinnati, OH	0.008	North Port, FL	0.003
Cleveland, OH	0.004	Ocala, FL	0.004
College Station, TX	0.052	Omaha, NE	0.004
Colorado Springs, CO	0.004	Orlando, FL	0.006
Columbia, MO	0.250	Ottawa, ON	0.003
Columbia, SC	0.002	Palm Bay, FL	0
Columbus, OH	0.004	Pensacola, FL	0.173
Dayton, OH	0.004	Philadelphia, PA	0.004
Deltona, FL	0.068	Pittsburgh, PA	0.003
Des Moines, IA	0.003	Port St. Lucie, FL	0.002
Durham, NC	0.003	Raleigh, NC	0.004
El Paso, TX	0.004	Richmond, VA	0.005
Fayetteville, AR	0.020	Riverside, CA	0.005
Fayetteville, NC	0.003	Roanoke, VA	0.005
Flagstaff, AZ	0.006	Salem, OR	0.003
Flint, MI	0.004	Salt Lake City, UT	0.005
Fort Collins, CO	0.003	Santa Fe, NM	0.004
Fresno, CA	0.004	Santa Rosa, CA	0.104
Gainesville, FL	0.005	South Bend, IN	0.002
Grand Rapids, MI	0.003	Spokane, WA	0.005
Green Bay, WI	0.003	Tacoma, WA	0.003
Greensboro, NC	0.003	Tallahassee, FL	0.003
Greenville, SC	0.003	Tampa, FL	0.004
Hartford, CT	0.003	Toledo, OH	0.003
Houston, TX	0.004	Toronto, ON	0.004
Jackson, MS	0.005	Tulsa, OK	0.004
Jacksonville, FL	0.004	Urban Honolulu, HI	0.004
Kahului, HI	0.003	Vancouver, WA	0.002
Kalamazoo, MI	0.003	Virginia Beach, VA	0.002
Kansas City, MO	0.004	Visalia, CA	0.004
Knoxville, TN	0.004	Waco, TX	0.005
Lafayette, IN	0.002	Wilmington, NC	0.007
Lakeland, FL	0.00368	Winston-Salem, NC	0.004
Lansing, MI	0.004	Worcester, MA	0.003

B Synthetic treatment

Consider $J+1$ units indexed j, where the first unit is untreated, and units $j = 2, ..., J+1$ are treated at time $t > T_0$, T_0 being the last pre-intervention period. Denote potential outcomes under treatment and non-treatment for unit j at time t as Y_{jt}^T and Y_{jt}^N, respectively. The estimand is $\alpha_{1t} = Y_{1t}^T - Y_{1t}^N$, for all t in which the donor group is treated.

Following Abadie (2021), we observe a vector of k outcome predictors for each unit, $X_{1j}, ..., X_{kj}$, where $\mathbf{X_1}, ..., \mathbf{X_{J+1}}$ collects the $k \times 1$ vectors of predictors for all j. Let $\mathbf{X_0}$ collect these vectors for units $j = 2, ..., J+1$. Finally, let $\mathbf{W}$ denote the $J \times 1$ vector of weights w_j for all treated units. The estimator of the potential outcome under treatment for the untreated unit is then

$$\hat{Y}_{1t}^T = \sum_{j=2}^{J+1} w_j Y_{jt}^T \tag{1}$$

and the treatment effect estimator is

$$\hat{\alpha}_{1t} = \hat{Y}_{1t}^T - Y_{1t}^N. \tag{2}$$

Taking as our predictor of interest the average value of the key outcome variable in the pre-treatment period, we choose $w_2, ..., w_{J+1}$ to minimize $\left| \bar{Y}_{1t}^N - \sum_{j=2}^{J+1} w_j \bar{Y}_{jt}^N \right|$. More generally, Abadie and Gardeazabal (2003) and Abadie et al. (2010) propose choosing $\mathbf{W}^*$ to minimize

$$(\mathbf{X_1} - \mathbf{X_0 W})' \mathbf{V} (\mathbf{X_1} - \mathbf{X_0 W}) \tag{3}$$

or the pre-treatment root mean squared prediction error (RMSPE), where $\mathbf{V}$ weights the relative importance of various predictors of the outcome in post-treatment. See Abadie (2021) for a deeper discussion of how to choose $\mathbf{V}$ in cases where more than one predictor is used.

Given estimated treatment effect $\hat{\alpha}_{1t}$, we follow Abadie et al. (2015) by constructing p-values using the sample analogue to

$$p = Pr \left(\left| \hat{\alpha}_{1t}^{PL} \right| \geq \left| \hat{\alpha}_{1t} \right| \right) \tag{4}$$

where the superscript *PL* indicates that we have run the synthetic treatment procedure on all potential donors and constructed an effect for each. The *p*-value will be large if, after running the synthetic treatment on every potential donor, we observe an effect that is frequently as large or larger for placebo units than for the untreated unit. We can also adjust *p*-values for pre-treatment fit by scaling the estimated treatment effect by the respective unit's pre-treatment RMSPE. Note that this inferential procedure can also be performed using the post-treatment RMSPE for the unit of interest and placebos, as a test of overall significance.

C   Meta-analysis details

We use Google Scholar to search for studies estimating the impact of ride-hailing on public transit ridership, traffic congestion, and traffic fatalities. We are interested in the net effect of ride-hailing on these three outcomes; papers that only estimate a partial effect are excluded (such as the effect of ride-hailing on alcohol-related traffic fatalities). We use the software Publish or Perish to find the first 500 articles on these topics found by Google Scholar on September 14, 2022.[32] To find other relevant studies, we inspect all the Google Scholar citations of papers that pass our first round of screening (again found using Publish or Perish).[33] This method is known as "snowballing."

We screen articles based on whether the title or abstract states or implies the article answers our research questions and whether the article is in English. We then read the paper to assess it more fully, requiring that the article must be empirical and include point estimates or bounds. We screen 2,596 articles, of which 56 are included in our dataset. A PRISMA diagram is provided in Figure A5.

Many papers report a range of estimates. For the meta-analysis, we focus on the estimates that the authors report as their main estimate. If they do not have a clear main estimate, then we take the median estimate.

In many cases we need to convert the estimates to make them comparable. If a

---

[32]Our search term is "('Uber' OR 'Lyft' OR 'Didi' OR 'TNC' OR 'transportation network company' OR 'ride-hail' OR 'ride-hailing' OR 'ridehail' OR 'ride-share' OR 'ride-sharing' OR 'rideshare') ('public transit' OR 'public transportation' OR 'congestion' OR 'speed' OR 'travel time' or 'delay' OR 'fatalities' OR 'deaths')."

[33]We also checked the references of all papers included in our final dataset, but all the articles we found were duplicates of those already found.

study reports results in levels, we convert the results to percentages by dividing by the mean. If a study estimated the effect using the natural log of the dependent variable, we convert to percentage changes using $\beta_\% = \exp \beta_{\log} - 1$. Traffic congestion is measured in a variety of ways, including travel speed, travel time, the ratio of the 95th percentile of travel times to free-flow travel time (known as the Planning Time Index), and vehicle kilometers traveled (VKT). While VKT is a measure of the amount of travel, rather than of congestion itself; increases in VKT are likely to lead to increased congestion.[34] For the sake of the meta-analysis, as much as possible, we convert estimates to be in terms of the impact on travel time.[35] When we have converted an estimate, we use the delta method to convert the standard error. We convert confidence intervals or t-statistics to standard errors by assuming the sampling distribution of the estimator is a normal distribution. For results that are statistically significant but only report the threshold (1%, 5%, 10%), we assume the standard error is the largest that still meets the threshold.[36]

For each article in our dataset, we record the geographic setting, time period, empirical method, and whether the article is published in a peer-reviewed journal. We also record for each outcome category, transit ridership, traffic congestion, traffic fatalities, the average effect and, if available, long-run effect, including the point estimate, standard error, sample size, and the time period after treatment. For transit ridership, we record estimates for total ridership, and separately by mode.

D   Publication bias

We next test for publication bias. A common way of testing for publication bias is using a funnel plot; however, this is less appealing in our setting as many articles in our sample do not include standard errors. However, when we test for an asymmetric funnel plot by regressing the absolute value of the point estimate

---

[34]Whether VKT leads to more congestion depends on when and where the increase in VKT occurs.

[35]If a paper reports VKT and vehicle hours traveled (VHT), but not speed, then we calculate average speed using average speed = VMT/VMH. We convert estimates of the percentage change in speed (β_s) to the percentage change in travel time (β_t) using $\beta_t = 1/(1 + \beta_s) - 1$.

[36]This assumption makes us more likely to find evidence of publication bias, as we are assuming that these studies barely passed the threshold for statistical significance.

on the standard error, we fail to reject the null hypothesis of no correlation between standard error and effect size (p-value 0.893, standard errors calculated by bootstrapping). This continues to hold when we include fixed effects for each outcome and whether an estimate is for the average (rather than long-run) effect (p-value 0.808).

We also test for publication bias by comparing the share of articles without statistically significant results (i.e., null results) that are published in peer-reviewed journals to the share of articles with statistically significant results. Column (1) of Table 13 shows that articles without a statistically significant average effect are actually more likely to be published in a peer-reviewed journal than those with a statistically significant main result. This difference is small and is not statistically significant. Columns (2)–(4) show that similar results hold within the subsamples focused on each of the three specific outcomes we study.

Given the well-known challenges of publishing null results, it is surprising to find so many published articles with null results. There are at least two explanations for what is happening in this literature. First, the outcomes we care about in this article are often not the primary focus of the articles included in the meta-analysis. For example, Young and Farber (2019) report on the "who, why, and when of Uber and other ride-hailing trips," which includes statistically insignificant results for the impact of ride-hailing on transit, but they also have many statistically significant results. Furthermore, six articles in our database include results on multiple outcomes we are interested in, such as having results on both transit ridership and traffic congestion (e.g., Li et al., 2021, Agarwal et al., 2023).[37]

Second, articles with a statistically insignificant estimate for the average effect of ride-hailing on one of our outcomes may find an effect in specific subgroups. Table 14 reports that articles where the average effect is not statistically significant are more than twice as likely to test for heterogeneous treatment effects than articles where the average effect is statistically significant. Among papers that estimate the effect of ride-hailing on public transit, those without a statistically significant average effect are three times as likely to test for heterogeneous treatment effects.

---

[37]While these six articles are more likely to have at least one of the main results not be statistically significant (50% vs. 32%), this difference is not statistically significant (Fisher's exact test p-value 0.397).

While these differences are not statistically significant at conventional levels, the p-values are remarkably small given the sample size.

This finding raises concerns about the "garden of forking paths" problem (Gelman and Loken, 2013). This problem is that researchers, without being deceitful or opportunistic, might alter their analyses based on the data they observe. One of the key examples Gelman and Loken (2013) give is "when the desired pattern does not show up as a main effect, it makes sense to look at interactions" (p. 4). For example, when a paper finds no effect of Uber on public transit, the authors start to look for (theoretically justified) heterogeneous treatment effects. Like Gelman and Loken (2013), we are not mentioning this to accuse researchers of malpractice; indeed, two of us have a paper in the meta-analysis that finds a null main effect but statistically significant interaction effects (Hall et al., 2018). But we agree with the recommendations in Gelman and Loken (2013) for observational studies like these and consider that these results are suggestive but not conclusive.

We have three caveats to this finding. First, articles with a null average effect often have a statistically significant long-run effect or an alternative specification that measures the intensity of treatment that is statistically significant. Second, our criteria for what counts as estimating heterogeneous treatment effects has some limitations. With the goal of capturing what the authors view as important, we only recorded heterogeneous treatment effects discussed in the title, abstract, introduction, or conclusion. Additionally, we do not consider estimating dynamic treatment effects nor estimating the effect on transit by mode to be heterogeneous treatment effects. Finally, some dimensions of heterogeneity have been fairly consistently tested within the literature, such as population (which also shows up as population density and urban compactness), which raises confidence that these findings are real.