Uber and Alcohol-Related Traffic Fatalities

Michael L. Anderson Lucas W. Davis*

July 2021

Abstract

Previous studies of the effect of ridesharing on alcohol-related traffic fatalities have yielded inconsistent, often contradictory conclusions. In this paper we revisit this question using proprietary data from Uber measuring monthly rideshare activity at the Census tract level. Most previous studies are based on publicly-available information about Uber entry dates into US cities, but we show that an indicator variable for whether Uber is available is a poor measure of rideshare activity — for example, it explains less than 3% of the tract-level variation in ridesharing, reflecting the enormous amount of variation both within and across cities. Using entry we find inconsistent and statistically insignificant estimates. However, when we use the more detailed proprietary data, we find a robust negative impact of ridesharing on alcohol-related traffic fatalities. Impacts concentrate during nights and weekends and are robust across a range of alternative specifications. Overall, our results imply that ridesharing has decreased US alcohol-related traffic fatalities by 6.1% and reduced total US traffic fatalities by 4.0%. Based on conventional estimates of the value of statistical life the annual life-saving benefits range from \$2.3 to \$5.4 billion. Back-of-the-envelope calculations suggest that these benefits may be of similar magnitude to producer surplus captured by Uber shareholders or consumer surplus captured by Uber riders.

Keywords: Drunk-driving, ridesharing, transportation network companies, value of statistical life
JEL: R41, R49, I12, I18

*(Anderson) Department of Agricultural and Resource Economics, University of California, Berkeley. Email: mlanderson@berkeley.edu. (Davis) Haas School of Business, University of California, Berkeley. Email: ldavis@haas.berkeley.edu. The authors have not received any financial compensation for this project nor do they have any financial relationships that relate to this research. They thank Luna Yue Huang for excellent research assistance, Jonathan D. Hall and Yizhen Gu for helpful comments, and Jonathan Wang, Santosh Rao Danda, and Cory Kendrick for assistance in accessing Uber data.

1 Background

There is a long history in economics of empirical analyses of traffic fatalities (Peltzman, 1975; Levitt and Porter, 2001b; Cohen and Einav, 2003; Ashenfelter and Greenstone, 2004; White, 2004; Anderson, 2008a; Abouk and Adams, 2013; Jacobsen, 2013; Anderson and Auffhammer, 2014; DeAngelo and Hansen, 2014). Analyses of alcohol and driving confirm that drunk drivers are an order of magnitude more dangerous than sober ones (Levitt and Porter, 2001a) and reveal that raising alcoholic beverage excise taxes and the minimum drinking age, and lowering blood-alcohol-content thresholds, are effective at reducing drunk-driving fatalities (Carpenter and Dobkin, 2009; Lovenheim and Slemrod, 2010; Sloan, 2020). More recently, a body of research has focused on the externalities of ridesharing. Ridesharing affects congestion (Hall et al., 2018; Tarduno, 2021), labor markets (Berger et al., 2018; Chen et al., 2019), and alcohol consumption (Teltser et al., 2021).

A number of recent studies estimate the effects of ridesharing on traffic fatalities. This existing work focuses on the timing of Uber entry into markets and yields inconsistent, often contradictory, conclusions. Brazil and Kirk (2016) and Zhou (2020) exploit the timing of Uber rollout across United States (US) counties and find no associations with traffic fatalities or drunk driving respectively. Dills and Mulholland (2018) find that the relationship between Uber entry and traffic fatalities can be negative or positive, depending on the choice of specification. Greenwood and Wattal (2017) and Peck (2017) focus on the timing of Uber rollout within California and New York City respectively; both studies find reductions in alcohol-related fatalities after the introduction of Uber. In contrast, Barrios et al. (2020) exploits city-level timing of Uber and Lyft rollout across the US and concludes that ridesharing *increases* traffic fatalities, while Cairncross et al. (2021) applies synthetic-control methods and finds a statistically insignificant relationship between ridesharing and traffic fatalities in Vancouver, BC.

In summary, the existing literature studies the effects of market entry by Uber and finds that it may cause traffic fatalities to decrease, increase, or remain unchanged.[1] Our study represents, to the best of our knowledge, the first work that uses proprietary Uber ridership data to estimate the effects of ridesharing on traffic fatalities. As we show in Section 3, market entry is a poor proxy for ridesharing activity, explaining less than 3% of the tract-level variation in ridesharing. When we emulate existing studies we find inconsistent and mostly statistically insignificant impacts. However, when we use the more detailed proprietary data, we find a robust negative impact of ridesharing on traffic fatalities. Impacts are negative and statistically significant across a range of alternative specifications and larger during nights and weekends, as expected.

[1] One exception is Morrison et al. (2017), which studies the disruption and resumption of Uber service in Las Vegas, Reno, San Antonio, and Portland. It finds mixed evidence — in Portland the resumption of Uber service correlated with a drop in alcohol-involved crashes, but in the other cities it did not.

The paper performs several back-of-the-envelope calculations aimed at putting our results in context. Scaled up to reflect current ridership levels, our results imply that ridesharing reduces total US alcohol-related traffic fatalities by 6.1% and reduces total US traffic fatalities by 4.0%. Based on conventional estimates of the value of statistical life (VSL), the annual life-savings benefits range from $2.3 to $5.4 billion. We compare these impacts to the market capitalization of Uber and estimates in the literature for the total consumer surplus from ridesharing.

2 Data

We combine data from two sources: the National Highway Traffic Safety Administration (NHTSA) Fatality Analysis Reporting System (FARS) and rideshare activity from Uber. FARS data represent a census of fatal crashes in the US. FARS contains detailed information on each crash, including geographic coordinates, roadway type, date and time, suspected involvement of alcohol, and driver characteristics. We include FARS data from 2001 to 2016.[2] Since Uber Technologies did not begin operations until 2010, and UberX (the service in which drivers typically use personal vehicles) did not launch until 2012, this long time series allows us to examine pre-trends in areas that subsequently experienced large increases in rideshare activity.

Appendix Table A1 reports summary statistics from the FARS data. In the US as a whole there are an average of 34,077 fatalities annually over the sample period. In areas that see rideshare activity by 2016 — our analytic sample — there are an average of 15,898 fatalities annually. In these areas, approximately 33% of fatalities involve multi-vehicle collisions, 40% involve single-vehicle collisions, and 26% involve pedestrians or bicyclists.

Alcohol involvement is reported in approximately 30% of fatal crashes. FARS contains several data elements pertaining to alcohol involvement. We use a variable that codes the number of drinking drivers involved in a crash. A driver qualifies as drinking if he has a positive blood alcohol concentration (BAC) or if police report alcohol involvement.[3] We classify the crash as involving alcohol if at least one driver was drinking. This necessarily undercounts the true number of alcohol-involved crashes because police may not always detect or report alcohol involvement, and alcohol data are "often missing", resulting in an "undercount [of] the actual number of drunk drivers" (National Highway Traffic Safety Administration 2016, p. 72). Thus, some fatal crashes with no reported alcohol involvement nevertheless involve alcohol, and we also estimate models that specify any fatal crash as the dependent variable.

Rideshare activity data come from Uber. These data report trip counts aggregated to the

[2]The 2000 file lacks crash geographic coordinates.

[3]According to the FARS codebook, a driver charged with an alcohol violation does not count as a drinking driver unless the driver also has a positive BAC or the police report alcohol involvement.

Census tract-by-month level for all Uber trips originating in a given Census tract, excluding those for which the origin or destination is an airport.[4] The data cover 1 July 2012 though 1 January 2017 for all Census tracts in the US, excluding those in Seattle, WA and New York City, NY (NYC).[5] Trip counts are normalized to the level of a specific Census tract in San Francisco, CA during May 2015, and Census tracts with less than 0.05% of this level of trip activity in a given month are rounded down to zero. We also observe data on the percentage of trips in a given tract-by-month observation that occurred between 8 PM on Friday and 4 AM on Sunday, a period during which drunk driving is more prevalent. While the authors had input into the data extraction parameters, decisions on what data to release ultimately rested with Uber staff.

To construct the analytic data set, we merged FARS data and Uber trip data at the tract-by-month level. Specifically, we assigned each crash in FARS to a Census tract and month using FARS data elements on longitude, latitude, and date. We then totaled different outcomes — e.g. total fatalities, total alcohol-involved fatalities, total fatalities by hour of day or day of week, etc. — at the tract-by-month level. Our largest estimation sample contains 2.7 million tract-by-month observations.

Uber trips are recorded based on Census tract of origin. The average Census tract, however, contains approximately 4,000 inhabitants, and a large county contains dozens of Census tracts. Many trips thus traverse more than one Census tract — the median Uber trip during our sample period was approximately five miles in length, and the 90th percentile trip was approximately ten miles in length.[6] To account for the multi-tract nature of most trips, we constructed the treatment of interest to be an inverse-distance-weighted average of Uber trip activity in nearby Census tracts. Specifically, for tract i we took a weighted average of trip activity in all tracts whose centroids are within 10 miles of tract i's centroid, with weights proportional to $distance_{ij}^{-1}$, where $distance_{ij}$ measures the distance in miles between tract i and tract j.[7] We confirm that our results are not sensitive to the exact

[4] Census tracts are based on 2010 Census definitions.

[5] To protect their proprietary data, Uber provided us with trip counts normalized to an arbitrary level. Public authorities in Seattle and NYC, however, publish detailed rideshare trip data. Thus, if our data set included Census tracts in these two cities, we would be able to back out raw trip counts.

[6] Source: Communications with Uber staff.

[7] Tract i's weight is normalized to one, and a tract whose centroid is one mile from tract i's centroid receives a weight of 0.25 since, in a perfect grid, each tract would be surrounded by four other tracts (note that the average tract size in a moderately-dense city is on the order of one square mile; for example, San José, CA is 181 square miles and contains approximately 200 Census tracts). In a simple gravity model in which trips radiate uniformly in all directions and continue indefinitely, the share of trips originating in j that cross tract i is proportional to the inverse of the distance between the two tracts. However, there are two complicating factors in our context. First, trips do not continue indefinitely; they die off with distance. Second, trips do not radiate uniformly, but tend to concentrate within travel corridors. The first factor suggests that the weight should decline more strongly with distance, while the second factor suggests that the weight should decline less strongly with distance. We assume that the two factors roughly offset each other, but we check that our results are robust to weights that decline somewhat more or less steeply than

specification of the weight, but given the spatially correlated nature of the treatment our analysis is best conceptualized as a city-level or metropolitan-area-level analysis.

One limitation of our data is that we do not measure activity for other ridesharing companies. During our sample period, however, Uber captured most of the US ridesharing market. For example, Lyft — Uber's largest US competitor by far — had approximately 6% and 14% of Uber's 2015 and 2016 US bookings respectively (the last two years of our sample; see Appendix A1 for details). Thus our data represent a good proxy for overall ridesharing activity during our sample period.

3 Empirical strategy and results

3.1 Estimates using publicly-available data

Before executing analyses with proprietary Uber data, we start by reporting results from a set of regressions based on publicly-available data. Following the approach used by several previous papers (see Section 2), we estimate regressions of the following form,

$$y_{it} = \beta_0 + \beta_1 1(\text{Rideshare Entry})_{it} + \delta_t + \gamma_i + \varepsilon_{it}. \tag{1}$$

The dependent variable y_{it} is an indicator for the presence of any fatal crash involving alcohol in tract i in month t, or, alternatively for any fatal crash of any kind.[8] When reporting results we multiply y_{it} by 100 for coefficient readability. The independent variable of interest, $1(\text{Rideshare Entry})_{it}$, is an indicator variable for whether Uber has entered the CBSA containing census tract i in month t.[9]

All specifications include month-of-sample (δ_t) and Census-tract (γ_i) fixed effects, controlling for any factors that change uniformly over time or differ across tracts in a time-invariant manner. The estimation sample runs from 1 January 2012 to 1 January 2017, and is restricted to include only tracts where Uber entry occured during this period. We exclude other census tracts because they tend to be less urban and unlikely to serve as a valid counterfactual.

Table 1 reports estimates and standard errors (clustered by state). Panels (A) and (B) report estimates for fatal crashes involving alcohol, and any fatal crash, respectively. Column (1) reports results for the complete sample of about 100 CBSAs, and then Columns

$distance_{ij}^{-1}$ (Appendix Tables A2 and A3).

[8] There is rarely more than one fatal crash in a tract-month observation, and there is typically only one fatality per fatal crash. Average total fatalities in a tract-month observation conditional on any fatal crash occurring are 1.13.

[9] To construct this measure we downloaded public data on Uber entry by city. Like most of the existing literature, we focus on UberX, the service that people generically refer to as "Uber". We matched cities to Core-Based Statistical Areas (CBSAs) and coded Census tracts accordingly. We performed this match for all cities with populations of 150,000 or more.

(2) through (4) restrict the sample to the first 50, 25, and 10 CBSAs where Uber entry occurred.

Consistent with the existing literature, the estimates are highly inconsistent, including some positive and some negative point estimates. Of the eight estimates, none are statistically significant. On the basis of the estimates in Table 1 it is impossible to draw strong conclusions about the relationship between ridesharing and traffic fatalities.

3.2 Main results

We now turn to analyses with proprietary data on Uber rideshare activity. We estimate a series of regressions identical to Equation (1), but we replace the indicator for Uber entry with our continuous measure of rideshare activity:

$$y_{it} = \beta_0 + \beta_1 \text{Rideshare}_{it} + \delta_t + \gamma_i + \varepsilon_{it} \tag{2}$$

In particular, the independent variable of interest, Rideshare_{it} is the rideshare activity index for tract i in month t, constructed as described in Section 2. We rescale the independent variable such that a value of 1 corresponds to the average value of the index in 2019 in our main analytic sample (a value that lies well within the support of our data).[10] The coefficient of interest thus approximates the average effect of ridesharing in 2019 on the probability of a fatal alcohol-involved crash, measured in percentage points.

As before, we limit the estimation sample to tracts that register nonzero ridesharing activity by the end of our sample; conceptually this strategy is similar to a "staggered adoption" design, but with a continuous treatment. We continue to include month-of-sample (δ_t) and Census-tract (γ_i) fixed effects. The identifying assumption is that, after controlling for tract and month fixed effects, ridesharing growth is uncorrelated with other tract-specific, time-varying factors that affect alcohol-related traffic fatalities. We probe this assumption in Section 3.4.

To account for dependence over time and across tracts, we cluster standard errors at the state level.[11] Since we explore a variety of specifications and samples in the paper, we also report false-discovery-rate (FDR) adjusted "q-values" (Benjamini and Hochberg, 1995). Briefly, the FDR represents the expected proportion of rejections that are false discoveries; controlling FDR at $q < 0.1$ thus indicates that 90% of rejections should be true rejections (Anderson, 2008b).[12]

[10] Uber activity in 2020 and 2021 was heavily impacted by COVID-19.

[11] Clustering at the county or CBSA level yields qualitatively similar results.

[12] For FDR control we define the family of hypotheses tested to include all main-body and appendix tables, with the exception of explicit placebo tests, which form their own family. The main family of hypotheses thus includes Tables 2, 3, 5, A2, A3, A4, A7, and Columns (5) and (6) of Table 4 (33 tests). The family of placebo hypotheses includes Columns (1) through (4) of Table 4 and Tables A5 and A6 (12 tests).

Table 2 reports estimates of β_1 for a range of sample restrictions. The first column estimates Equation (2) using all tracts with nonzero ridesharing activity by January 2017 (approximately 45,000 tracts). A one unit increase in ridesharing activity reduces the probability of an alcohol-related fatal crash by 0.038 percentage points ($t = -3.2$). This corresponds to approximately a 4.8% decrease in alcohol-related fatalities. Columns (2) through (5) estimate Equation (2) when restricting the sample to include tracts whose endpoint rideshare activity falls within the top 50, 25, 15, and 10 percent of all tracts respectively. The estimate of β_1 in Column (2) implies that a one unit increase in ridesharing activity reduces alcohol-related fatalities by 0.043 percentage points ($t = -3.6$), or 6.1% of the mean. The estimates of β_1 in Columns (3) and (4) are similar in magnitude and remain highly significant. In Column (5) the implied effect (-0.03 percentage points) is 6% of the mean and remains statistically significant, despite dropping 90% of tracts.

The evidence in Table 2 implies a significant negative relationship between ridesharing and alcohol-related fatalities.[13] The consistent effects in Table 2 stand in contrast to the varied conclusions in Table 1 and in the existing previous literature. The results suggest that rideshare entry is a poor proxy for true rideshare activity during our sample period; indeed, a regression of rideshare activity on rideshare entry yields a R^2 of 0.03. The indicator for Uber entry fails to capture the large variation in the intensity of ridesharing both within and across CBSAs.

3.3 Nights and weekends

We next focus on effects during nights and weekends. We expect the effects of ridesharing activity to concentrate during nights and weekends for two reasons. First, drunk-driving fatalities concentrate during nights and weekends; in our analytic sample, approximately 75 percent of alcohol-involved fatalities occur between 8 pm and 6 am (see Appendix Figures A2 and A3). Second, daytime drunk-driving fatalities are less plausibly related to dining and entertainment trips, and more plausibly related to alcohol abuse. To the extent that ridesharing is a better substitute for driving for dining and entertainment trips, ridesharing activity should have stronger effects during night than day.

Table 3 reports estimates of Equation (2) when limiting measurement of the dependent or independent variables to nights and weekends. We use the top 50 percent sample from Column (2) of Table 2 for the analytic sample. Column (1) reproduces the top 50 percent estimate from Table 2 for comparison purposes. Column (2) reports $\hat{\beta}_1$ when Rideshare_{it} is computed using only rides occurring between 8 pm Friday and 4 am Sunday. The estimate

[13] Appendix Tables A2 and A3 reproduce Table 2 using an independent variable that applies a weight proportional to $distance^{-0.9}$ or $distance^{-1.1}$ respectively (as compared to $distance^{-1}$) The statistical significance and coefficient magnitudes are broadly similar to Table 2. If anything the patterns in Appendix Table A2 suggest that using a distance weight that declines less steeply might yield somewhat larger estimates. To remove researcher discretion, however, we prefer the simple inverse distance weight.

of β_1 is of similar magnitude and remains highly significant, suggesting that weekend ridesharing activity alone is a sufficient statistic for predicting effects on alcohol-involved fatalities.[14] Column (3) reports $\hat{\beta}_1$ when y_{it} is computed using only fatalities occurring between 8 pm and 6 am. The estimate of β_1 changes to -0.026 and remains significant. Column (4) combines both restrictions from Columns (2) and (3), while Columns (5) and (6) further limit y_{it} to fatalities occurring between 8 pm Friday and 4 am Sunday. In all three columns the estimates of β_1 remain negative, though they lose significance in Columns (4) and (6).

3.4 Threats to identification

We anticipate two primary threats to identification in Equation (2): differential trends in factors affecting drunk driving and differential trends in factors affecting traffic fatality rates. Examples of the first threat include alcohol prices, drunk-driving penalties, and enforcement. Examples of the second threat include changes in vehicle miles traveled, infrastructure improvements, vehicle fleet technology, driver characteristics, and health care systems.

To address the first threat, we modify Equation (2) to include state-by-month fixed effects, δ_{st}:

$$y_{ist} = \beta_0 + \beta_1 \text{Rideshare}_{ist} + \delta_{st} + \gamma_i + \varepsilon_{ist} \tag{3}$$

Ridesharing growth in our sample occurs over a period of only several years, and most policies that could quickly affect alcohol prices or drunk driving — including alcohol taxes, drunk-driving penalties, and enforcement by state highway patrols — vary primarily at the state level. State-by-month fixed effects thus absorb many potential confounders of interest. Appendix Table A4 reports results from estimating Equation (3) with all alcohol-related fatalities (Columns (1) and (3)) or nighttime alcohol-related fatalities (Column (2)) as the dependent variable. The first two columns limit the sample to tracts with above median endpoint rideshare activity, and the last column limits the sample to tracts in the top 25 percent of endpoint rideshare activity. In all three columns the estimates of β_1 are of similar magnitude to the analogous estimates from Tables 2 and 3 and they remain highly significant.

To address the second threat, we use daylight hours to execute a "triple-differences" design. Most factors that affect traffic fatality rates should impact both daytime and nighttime fatalities. Testing for effects during daytime hours thus represents an important falsification test for our research design. Table 4 reports estimates of Equation (3) when limiting

[14] As in Column (1), we rescale the independent variable in Column (2) such that a value of 1 corresponds to the approximate average value of the index in 2019. The coefficient in this column does not have a direct causal interpretation, however, since non-weekend ridesharing activity correlates with weekend ridesharing activity.

measurement of the dependent variable to daytime hours (8 am to 5 pm). Columns (1) and (2) report estimates of β_1 for tracts whose endpoint rideshare activity is within the top 50 percent or the top 25 percent respectively. In both cases $\hat{\beta}_1$ is close to zero, precisely estimated, and statistically insignificant.

Falsification tests using drunk-driving fatalities, however, may be limited in what they can detect because the majority of alcohol-related fatalities occur during nighttime. Columns (3) and (4) reproduce Columns (1) and (2) but replace y_{it} with an indicator for *any* fatal crash in tract i and month t during daytime hours. The estimates of β_1 are close to zero, precisely estimated, and statistically insignificant. These results imply the absence of differential trends in traffic fatalities in high rideshare growth tracts.

To formally estimate the triple-differences design we specify the following regression:

$$y_{itd} = \beta_0 + \beta_1 \text{Rideshare}_{it} \cdot \mathbf{1}(d = 1) + \delta_{td} + \gamma_{it} + \phi_{id} + \alpha_{std} + \varepsilon_{itd} \tag{4}$$

where y_{itd} is an indicator for any alcohol-related fatalities in tract i in month t during time-of-day d. We let $d = 0$ correspond to hours between 8 am and 5 pm (daytime), and $d = 1$ correspond to all other hours. The coefficient of interest, β_1, is on the interaction between Rideshare_{it} and an indicator for non-daytime hours. To operationalize the triple-differences design Equation (4) includes month-by-time-of-day (δ_{td}), tract-by-month (γ_{it}), and tract-by-time-of-day (ϕ_{id}) effects, absorbing all but the *itd*-level variation. For completeness we also include the state-by-month fixed effects, now interacted with time-of-day (α_{std}).

Columns (5) and (6) report estimates of β_1 from Equation (4), using the same estimation samples as Columns (1) and (2) respectively. In both cases $\hat{\beta}_1$ is similar in magnitude to the corresponding estimate from Table 2 (Columns (2) or (3)) and statistically significant, as we might expect given the null effects in Columns (1) and (2).

As a final challenge to our research design, we reproduce estimates of Equation (2) when shifting the treatment back in time by seven (2005 to 2009) or twelve years (2001 to 2004). We use these two nonoverlapping periods because they predate the launch of the Uber app.[15] These placebo estimates, presented in the first two columns of Appendix Table A5, test whether future changes in ridesharing activity predict current changes in alcohol-related fatalities. The estimates of β_1 are statistically insignificant, suggesting that tracts with high future ridesharing growth did not have different alcohol-related fatality trends in the years prior to the launch of Uber than those with lower future ridesharing growth.

3.5 All traffic fatalities

While alcohol-related fatalities are the focus of our analysis, rideshare activity may affect non-alcohol-related traffic fatalities as well for two reasons. First, some fatal crashes without reported alcohol involvement may nevertheless involve alcohol, as noted in Section 2.

[15] Uber launched the "UberCab" app in San Francisco in 2010.

For example, even having a driver charged with an alcohol violation does not qualify a crash as involving alcohol, unless the driver also has a positive BAC or the police report alcohol involvement. Second, ridesharing drivers may be safer (or less safe) than the drivers they replace, regardless of drunk-driving behavior.[16]

To measure the effect of rideshare activity on all traffic fatalities, we estimate versions of Equations (3) and (4) that specify the dependent variable as an indicator for any fatal crash. Columns (1) and (2) in Table 5 estimate Equation (3) with y_{it} specified as an indicator for any fatal crash. The estimates of β_1 are highly significant and roughly twice as large as the analogous estimates in Table 2 (Columns (2) and (3)), suggesting that some fatal crashes without reported alcohol involvement nevertheless involve alcohol. Columns (3) and (4) in Table 5 estimate Equation (4), the triple-differences design, with y_{itd} specified as an indicator for any fatal crash. The estimates of β_1 are of approximately similar magnitude as in Columns (1) and (2) and remain highly significant.[17] Though the point estimates for effects on any fatalities are larger than those for alcohol-involved fatalities, it is worth clarifying that the implied percentage effects are substantially smaller for total fatalities than for alcohol-related fatalities, as is evident from comparing the point estimates to the respective means of the dependent variables.

4 Discussion

The coefficient estimates in Table 2 represent the monthly effect of the average level of 2019 Uber activity in tracts contained in our main analytic sample (Column (2) of Table 2). To compute annual alcohol-related fatal crashes avoided due to Uber circa 2019, we multiply these coefficients by the number of sample tracts (36,780) times 12.[18] Finally, we multiply by the average number of fatalities conditional on any fatal crash occurring (1.13).

Our estimates of the effects on alcohol-related fatalities imply that Uber saved 214 lives in 2019, or a reduction of approximately 6.1%. We compute this estimate using Column (2) of Table 2, but using the coefficient from Column (3), for example, generates broadly

[16] Ridesharing activity could also affect total travel. However, we expect that the increase in travel would be weakly positive, which would generally increase traffic fatalities.

[17] As in Section 3.4, we estimate placebo specifications that shift the treatment back in time by seven or twelve years. These placebo estimates, presented in Appendix Table A6, test whether future changes in ridesharing activity predict current changes in traffic fatalities. The estimates of β_1 are statistically insignificant in three of four columns. The estimate is at the margin of significance in Column (2), but insignificant after FDR adjustment. Overall these results reveal that tracts with high future ridesharing growth did not consistently have different fatality trends in the years prior to the launch of Uber than those with lower future ridesharing growth, but they also suggest that the models with all traffic fatalities may be slightly less robust than those with alcohol-related fatalities only.

[18] Multiplying by 12 annualizes the estimates, since our data are at the tract-by-month level. We also divide by 100 since the tables multiply y_{it} by 100 for coefficient readability.

similar estimates.[19] Alternatively, we can compute the effects of Uber on total fatalities using the estimate from Column (1) of Table 5. This coefficient implies that Uber saved 494 lives in 2019, or a reduction of approximately 4.0%. Note that these calculations include lives saved by Uber only; total lives saved by ridesharing would also include the impacts of competitors such as Lyft.

To understand the economic magnitudes of these estimates, we apply the Department of Transportation value of a statistical life (VSL) of $10.9 million ($2019, US Department of Transportation 2021). The annual life-saving benefits range from $2.3 billion (214 lives saved) to $5.4 billion (494 lives saved). These benefits represent a mixture of internally and externally captured benefits. To approximate the share of benefits that are internal versus external, we estimate Equation (2) separately for single-vehicle crashes, multi-vehicle crashes, and pedestrian/bicyclist-involved crashes. These results, reported in Appendix Table A7 suggest that single-vehicle, multi-vehicle, and pedestrian crashes account for approximately 30%, 54%, and 16% of fatalities involving drunk drivers respectively. If half the multi-vehicle fatalities and all of the pedestrian/bicyclist fatalities are external, then 43% of the life-saving benefits represent pure externalities.[20]

Of the remaining 57% of life-saving benefits, an indeterminate fraction are internalized. While drunk individuals by definition have impaired decision-making skills, they may choose whether to drive or rideshare at a point when they are still sober. If they fully understand the risks of drunk driving at that point, they should internalize their own safety benefits of ridesharing. Nevertheless, the general consensus in the economic literature is that the decision to drink and drive is not fully rational (Sloan, 2020); for example, drinker-drivers exhibit time-inconsistent preferences (Sloan et al., 2014). We conservatively assume that drinking drivers understand and internalize the vast majority (7/8ths) of their private safety benefits. Under that assumption, half of the total life-saving benefits represent external benefits ($1.2 to $2.7 billion), and the other half are internalized by riders.

We compare these external and internal benefits against two benchmarks: producer surplus captured by Uber shareholders and consumer surplus captured by Uber riders. Uber's

[19] To compute total lives saved using only the top 25 percent of tracts, i.e. Column (3), we first normalize the index by the average level of Uber activity in these tracts. Doing so transforms the coefficient in Column (3) to -0.076. We then perform the same calculation as above and find that, in just the top 25 percent of tracts, Uber saved approximately 189 lives in 2019. Intuitively, the majority of lives saved are concentrated in the areas with the highest Uber ridership.

[20] In principle, tort liability rules and insurance mandates could internalize the externality in these cases. Nevertheless, even if the drunk driver is found liable for the crash, almost no drivers possess assets that are sufficient to cover a VSL of $10 million. Furthermore, state-mandated levels of liability insurance are grossly inadequate to cover the costs of a fatality; only Maine require drivers to carry more than $30,000 of liability coverage for each person injured (Insurance Information Institute, 2021). Many drivers remain uninsured despite the regulations, and few drivers have policies that exceed several hundred thousand dollars of coverage.

market capitalization represents producer surplus captured by Uber shareholders. In 2019, Uber's average market capitalization was \$51.3 billion. After adjusting for the proportion of revenues outside the US or not involving ridesharing, we estimate that domestic ridesharing accounted for \$19.2 billion of this market capitalization.[21] To convert from stocks to flows we apply a price-earnings (PE) ratio of 22.7, the PE ratio for the S&P 500 during this period.[22] We thus calculate that US ridesharing producer surplus captured by Uber shareholders was equivalent to an annual stream of \$0.9 billion (\$19.2b / 22.7). The external life-saving benefits of Uber (\$1.2 to \$2.7 billion) — which represent pure unmeasured welfare gains — thus exceed shareholder producer surplus.

Estimating consumer surplus accruing from Uber's existence is challenging because it depends on the long-run price elasticity of demand. A recent experiment by Christensen and Osman (2021) estimates the medium-run price elasticity of demand for Uber by randomly assigning large discounts of 25% to 50% to Uber users for three months. While this paper has limitations in our context — e.g. the three-month treatment is not truly long-run, and it was conducted in Cairo — it is to our knowledge the best available evidence on this parameter.[23] Christensen and Osman (2021) estimates a price elasticity of approximately -8.

A long-run price elasticity of -8 implies that consumer surplus equals one-seventh of total spending (Appendix A2).[24] Total spending by domestic Uber riders in 2019 was approximately \$24.7 billion.[25] An estimate of 2019 consumer surplus is thus $\$24.7\text{b}/7 =$

[21]The US and Canada accounted for 62% of 2019 revenue, and ridesharing accounted for 75% of 2019 revenue (Uber Technologies 2020, p. 61 and p. 115). We assume that the US accounted for 90% of US and Canada revenue, based on their respective populations. Finally, we net out NYC and Seattle revenue using publicly available ridership figures, as those cities are not in our data. See Appendix A1 for details.

[22]This ratio is likely conservative since Uber is a "growth" stock.

[23]Cohen et al. (2016) estimates the short-run price elasticity of demand for Uber using a clever natural experiment generated by surge pricing. The short-run price elasticity of demand is not our object of interest, however; as Cohen et al. (2016) point out, "If...one wanted to know how consumers would be affected if Uber disappeared permanently, a long-run elasticity would be more appropriate." (p. 21) Similarly, Goldszmidt et al. (2020) estimate the value-of-time using a field experiment with Lyft. As part of this experiment they randomly vary the surge-pricing multiplier to identify price elasticities. While the manipulation can last up to eight weeks, it applies only to surge pricing, and all changes in quantities are measured conditional on launching the app (i.e. the relevant price variation is likely the difference between the price a user expects when launching the app and the actual displayed price). Thus the price elasticity measured is primarily short-run in nature, which works well for the authors' purposes but is less relevant to our welfare calculations. The results in both papers suggest that the short-run price elasticity of demand is likely inelastic (less than one in absolute value).

[24]Notably, this estimate is not far outside the range reported in Shapiro (2018), which estimates a structural model and concludes that consumer surplus from Uber in NYC ranges from approximately 2% of fares in central Manhattan to 10% of fares in less dense outer boroughs.

[25]Uber gross ridesharing bookings were \$49.7 billion in 2019 (Uber Technologies 2020, p. 61). To compute US ridesharing bookings we apply the same country-specific shares as in Footnote 21 and net out NYC and Seattle. See Appendix A1 for details.

\$3.5 billion. The internalized life-saving benefits alone would thus represent between 34% and 77% of total consumer surplus. At a price elasticity of -4 the internalized life-saving benefits represent between 15% and 33% of consumer surplus, and at a price elasticity of -2, the internalized life-saving benefits still represent between 6% and 14% of consumer surplus.

5 Conclusion

Previous researchers have worked hard to learn as much as possible from publicly-available information about Uber. Our analysis suggests, however, that whether Uber is operating in a given metropolitan area is an inherently poor proxy for ridesharing activity. When we instead use proprietary tract-level information on Uber ridership, the impacts come into sharper focus, and we find robust, large, and statistically significant negative impacts on alcohol-related traffic fatalities.

Last year in the United States there were 42,000 traffic fatalities, including over 10,000 that were alcohol-related. The total economic damages, applying a standard VSL, approach half a *trillion* dollars. Understanding the factors that contribute to these deaths continues to be an important question for economists and other researchers. Our results suggest that ridesharing can play an important role in reducing these deaths, and that these benefits may represent a meaningful fraction of the total consumer surplus from ridesharing.

References

Abouk, Rahi and Scott Adams, "Texting bans and fatal accidents on roadways: do they work? Or do drivers just react to announcements of bans?," *American Economic Journal: Applied Economics*, 2013, *5* (2), 179–99.

Anderson, Michael, "Safety for Whom? The Effects of Light Trucks on Traffic Fatalities," *Journal of Health Economics*, 2008, *27* (4), 973–989.

Anderson, Michael L, "Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects," *Journal of the American Statistical Association*, 2008, *103* (484), 1481–1495.

_ **and Maximilian Auffhammer**, "Pounds that Kill: The External Costs of Vehicle Weight," *Review of Economic Studies*, 2014, *81* (2), 535–571.

Ashenfelter, Orley and Michael Greenstone, "Using Mandated Speed Limits to Measure the Value of a Statistical Life," *Journal of Political Economy*, 2004, *112* (S1), S226–S267.

Barrios, John M, Yael Hochberg, and Hanyi Yi, "The cost of convenience: Ride-hailing and traffic fatalities," 2020. National Bureau of Economic Research Working Paper.

Benjamini, Y. and Y. Hochberg, "Controlling the False Discovery Rate," *Journal of the Royal Statistical Society: Series B*, 1995, *57*, 289–300.

Berger, Thor, Chinchih Chen, and Carl Benedikt Frey, "Drivers of disruption? Estimating the Uber effect," *European Economic Review*, 2018, *110*, 197–210.

Brazil, Noli and David S Kirk, "Uber and Metropolitan Traffic Fatalities in the United States," *American Journal of Epidemiology*, 2016, *184*, 192–198.

Cairncross, John, Jonathan D. Hall, and Craig Palsson, "VancUber: The effect of ride-hailing on public transportation, congestion, and traffic fatalities," 2021. University of Toronto Working Paper.

Carpenter, Christopher and Carlos Dobkin, "The effect of alcohol consumption on mortality: regression discontinuity evidence from the minimum drinking age," *American Economic Journal: Applied Economics*, 2009, *1* (1), 164–82.

Chen, M Keith, Peter E Rossi, Judith A Chevalier, and Emily Oehlsen, "The value of flexible work: Evidence from uber drivers," *Journal of Political Economy*, 2019, *127* (6), 2735–2794.

Christensen, Peter and Adam Osman, "The Demand for Mobility: Evidence from an Experiment with Uber Riders," 2021. National Bureau of Economic Research Working Paper.

Cohen, Alma and Liran Einav, "The Effects of Mandatory Seat Belt Laws on Driving Behavior and Traffic Fatalities," *Review of Economics and Statistics*, 2003, *85* (4), 828–843.

Cohen, Peter, Robert Hahn, Jonathan Hall, Steven Levitt, and Robert Metcalfe, "Using Big Data to Estimate Consumer Surplus: The Case of Uber," 2016. National Bureau of Economic Research Working Paper.

DeAngelo, Gregory and Benjamin Hansen, "Life and Death in the Fast Lane: Police Enforcement and Traffic Fatalities," *American Economic Journal: Economic Policy*, 2014, *6* (2), 231–257.

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

Goldszmidt, Ariel, John A. List, Robert D. Metcalfe, Ian Muir, V. Kerry Smith, and Jenny Wang, "The Value of Time in the United States: Estimates from Nationwide Natural Field Experiments," 2020. National Bureau of Economic Research Working Paper.

Greenwood, Brad N and Sunil Wattal, "Show Me the Way to Go Home: An Empirical Investigation of Ride-Sharing and Alcohol Related Motor Vehicle Fatalities," *MIS Quarterly*, 2017, *41* (1).

Hall, Jonathan D., Craig Palsson, and Joseph Price, "Is Uber a substitute or complement for public transit?," *Journal of Urban Economics*, 2018, *108*, 36–50.

Insurance Information Institute, "Background on: Compulsory Auto/Uninsured Motorists," 2021.

Jacobsen, Mark R, "Fuel Economy and Safety: The Influences of Vehicle Class and Driver Behavior," *American Economic Journal: Applied Economics*, 2013, *5* (3), 1–26.

Levitt, Steven D and Jack Porter, "How Dangerous are Drinking Drivers?," *Journal of Political Economy*, 2001, *109* (6), 1198–1237.

_ **and** _ , "Sample Selection in the Estimation of Air Bag and Seat Belt Effectiveness," *Review of Economics and Statistics*, 2001, *83* (4), 603–615.

Lovenheim, Michael F and Joel Slemrod, "The fatal toll of driving to drink: The effect of minimum legal drinking age evasion on traffic fatalities," *Journal of Health Economics*, 2010, *29* (1), 62–77.

Lyft, "Form S-1 Registration Statement," March 2019.

Morrison, Christopher N, Sara F Jacoby, Beidi Dong, M Kit Delgado, and Douglas J Wiebe, "Ridesharing and Motor Vehicle Crashes in 4 US Cities: An Interrupted Time-Series Analysis," *American Journal of Epidemiology*, 06 2017, *187* (2), 224–232.

National Highway Traffic Safety Administration, "Fatality Analysis Reporting System (FARS) Analytical User's Manual 1975-2015," 2016.

Peck, Jessica Lynn, "New York City Drunk Driving After Uber," 2017. CUNY Working Paper.

Peltzman, Sam, "The Effects of Automobile Safety Regulation," *Journal of Political Economy*, 1975, *83* (4), 677–725.

Shapiro, Matthew, "Density of demand and the benefit of Uber," 2018. Singapore Management University Working Paper.

Sloan, Frank A., "Drinking and Driving," *Handbook of Labor, Human Resources and Population Economics*, 2020, pp. 1–31.

Sloan, Frank A, Lindsey M Eldred, and Yanzhi Xu, "The behavioral economics of drunk driving.," *Journal of Health Economics*, May 2014, *35*, 64–81.

Tarduno, Matthew, "The congestion costs of Uber and Lyft," *Journal of Urban Economics*, 2021, *122*, 103318.

Teltser, Keith, Conor Lennon, and Jacob Burgdorf, "Do ridesharing services increase alcohol consumption?," *Journal of Health Economics*, 2021, *77*, 102451.

Uber Technologies, "Form S-1 Registration Statement," April 2019.

_ , "2019 Annual Report," 2020.

US Department of Transportation, "VSL Update 2021 - Transmittal Memo," 2021.

White, Michelle J., "The "Arms Race" on American Roads: The Effect of Sport Utility Vehicles and Pickup Trucks on Traffic Safety," *The Journal of Law and Economics*, Oct 2004, *47* (2), 333–355.

Zhou, You, "Ride-sharing, alcohol consumption, and drunk driving," *Regional Science and Urban Economics*, 2020, *85*, 103594.

Table 1: Emulating Existing Studies Using Data on Uber Entry

	(1)	(2)	(3)	(4)
	A. Dependent variable: Drunk death			
Rideshare entry	-0.020	-0.027	-0.039	-0.033
	(0.020)	(0.024)	(0.026)	(0.022)
Dependent variable mean	0.798	0.738	0.577	0.491
	B. Dependent variable: Any death			
Rideshare entry	0.018	-0.014	-0.040	0.020
	(0.036)	(0.041)	(0.063)	(0.051)
Observations	2,774,640	2,067,420	1,364,700	812,400
Tracts	46,244	34,457	22,745	13,540
Dependent variable mean	2.791	2.567	2.154	1.859
CBSAs	Any entry	First 50	First 25	First 10

Notes: This table reports coefficient estimates from eight separate least squares regressions. The unit of observation is the Census tract by month. The dependent variable in all regressions is an indicator for alcohol-related fatalities (Panel A) or any fatalities (Panel B), multiplied by 100. The independent variable is an indicator for whether Uber has entered the CBSA containing Census tract i in month t. All regressions include Census-tract and month-of-sample fixed effects. Parentheses contain standard errors clustered by state.

Table 2: The Effect of Uber on Alcohol-Related Traffic Fatalities, Main Results

	(1) Drunk death	(2) Drunk death	(3) Drunk death	(4) Drunk death	(5) Drunk death
Rideshare index	-0.038	-0.043	-0.040	-0.039	-0.030
	(0.012)	(0.012)	(0.013)	(0.014)	(0.012)
	0.007	*0.003*	*0.006*	*0.009*	*0.016*
Observations	2,688,480	2,206,800	1,098,840	657,900	438,840
Max tract rideshare activity	Nonzero	Top 50 pct	Top 25 pct	Top 15 pct	Top 10 pct
Tracts	44,808	36,780	18,314	10,965	7,314
Mean dependent variable	0.788	0.705	0.609	0.556	0.497
Mean index	0.119	0.145	0.280	0.437	0.605

Notes: This table reports coefficient estimates from five separate least squares regressions that progressively restrict the sample to locations with higher ridesharing activity by the end of our sample period. The unit of observation is the Census tract by month. The dependent variable in all regressions is an indicator for alcohol-related fatalities, multiplied by 100. The independent variable is the weighted average of rideshare activity originating within a 10-mile radius of Census tract i in month t, with weights equal to $0.25 \cdot \text{distance}^{-1}$, normalized such that a value of 1 corresponds to the average value of the index in 2019. All regressions include Census-tract and month-of-sample fixed effects. Parentheses contain standard errors clustered by state. FDR-control q-values in *italics*.

Table 3: The Effect of Uber on Alcohol-Related Traffic Fatalities, Nights and Weekends

	(1) Drunk death	(2) Drunk death	(3) Drunk death	(4) Drunk death	(5) Drunk death	(6) Drunk death
Rideshare index	-0.043	-0.039	-0.026	-0.022	-0.017	-0.014
	(0.012)	(0.012)	(0.013)	(0.013)	(0.008)	(0.009)
	0.003	*0.006*	*0.046*	*0.105*	*0.044*	*0.114*
Observations	2,206,800	2,206,800	2,206,800	2,206,800	2,206,800	2,206,800
Dependent variable coverage	All	All	8p-6a	8p-6a	Fr 8p-Su 4a	Fr 8p-Su 4a
Independent variable coverage	All	Fr 8p-Su 4a	All	Fr 8p-Su 4a	All	Fr 8p-Su 4a
Tracts	36,780	36,780	36,780	36,780	36,780	36,780
Mean dependent variable	0.705	0.705	0.506	0.506	0.288	0.288
Mean index	0.145	0.160	0.145	0.160	0.145	0.160

Notes: This table reports coefficient estimates from six separate least squares regressions that progressively restrict the hours used to measure the dependent variable, independent variable, or both, to nights and weekends. The unit of observation is the Census tract by month. The sample includes only tracts with rideshare activity in the top 50% by the end of our sample period. The dependent variable in all regressions is an indicator for alcohol-related fatalities, multiplied by 100. The independent variable is the weighted average of rideshare activity originating within a 10-mile radius of Census tract i in month t, with weights equal to $0.25 \cdot \text{distance}^{-1}$, normalized such that a value of 1 corresponds to the average value of the index in 2019. All regressions include Census-tract and month-of-sample fixed effects. Parentheses contain standard errors clustered by state. FDR-control q-values in *italics*.

Table 4: The Effect of Uber on Alcohol-Related Traffic Fatalities, Daytime Hours and Triple Differences

	(1) Drunk death	(2) Drunk death	(3) Any death	(4) Any death	(5) Drunk death	(6) Drunk death
Rideshare index	-0.005	0.001	-0.010	-0.001		
	(0.003)	(0.002)	(0.015)	(0.011)		
	0.343	*0.787*	*0.787*	*0.951*		
Rideshare index * Night					-0.038	-0.042
					(0.016)	(0.011)
					0.021	*0.002*
Observations	2,206,800	1,098,780	2,206,800	1,098,780	4,413,600	2,197,560
Max tract rideshare activity	Top 50 pct	Top 25 pct	Top 50 pct	Top 25 pct	Top 50 pct	Top 25 pct
Tracts	36,780	18,314	36,780	18,314	36,780	18,314
Mean dependent variable	0.0829	0.0626	0.734	0.618	0.352	0.304
Mean index	0.145	0.280	0.145	0.280	0.0726	0.140

Notes: This table reports coefficient estimates from different least squares regressions. The unit of observation is the Census tract by month (Columns (1) through (4)) or Census tract by month by time-of-day (Columns (5) and (6)). The dependent variable in Columns (1) and (2) is an indicator for alcohol-related fatalities during daytime hours (8 am to 5 pm), multiplied by 100. The dependent variable in Columns (3) and (4) is an indicator for any fatalities during daytime hours, multiplied by 100. Finally, the dependent variable in Columns (5) and (6) is an indicator for alcohol-related fatalities in either daytime or non-daytime hours, multiplied by 100. In all regressions, the independent variable is the weighted average of rideshare activity originating within a 10-mile radius of Census tract i in month t, with weights equal to $0.25 \cdot \text{distance}^{-1}$, normalized such that a value of 1 corresponds to the average value of the index in 2019. All regressions include Census-tract and state-by-month-of-sample fixed effects; Columns (5) and (6) include tract-by-month-of-sample, time-of-day-by-month-of-sample, tract-by-time-of-day, and state-by-month-of-sample-by-time-of-day fixed effects. Parentheses contain standard errors clustered by state. FDR-control q-values in *italics*.

Table 5: The Effect of Uber on All Traffic Fatalities, State-by-Month FEs/Triple Differences

	(1) Any death	(2) Any death	(3) Any death	(4) Any death
Rideshare index	-0.099	-0.089		
	(0.021)	(0.018)		
	0.001	*0.001*		
Rideshare index * Night			-0.078	-0.088
			(0.022)	(0.015)
			0.003	*0.001*
Observations	2,206,800	1,098,780	4,413,600	2,197,560
Max tract rideshare activity	Top 50 pct	Top 25 pct	Top 50 pct	Top 25 pct
Tracts	36,780	18,314	36,780	18,314
Mean dependent variable	2.458	2.242	1.229	1.121
Mean index	0.145	0.280	0.0726	0.140

Notes: This table reports coefficient estimates from four separate least squares regressions. Specifications are identical to Table 2, Columns (2) and (3) (with the inclusion of state-by-month-of-sample fixed effects), and Table 4, Columns (5) and (6), except that the dependent variable is an indicator for any fatalities (rather than alcohol-related fatalities), multiplied by 100. Parentheses contain standard errors clustered by state. FDR-control q-values in *italics*.

Appendix

Not For Print Publication

A1 Ridesharing revenue and market capitalization

A1.1 Early Lyft market share

To compute Lyft marketshare in 2015 and 2016, we first estimate Lyft gross bookings (i.e. rider payments). Lyft (2019) reports 2016 revenue as $343m (p. 86), and we estimate 2015 revenue as $72m based on reported 2016 ridership growth of 575% (p. 82). 2016 revenue as a percentage of bookings was 18% (p. 82), yielding gross bookings of $1.9b and $400m in 2016 and 2015 respectively. Note that Lyft operated exclusively in the US and Canada during these years.

Uber Technologies (2019) reports 2016 US and Canada ridesharing revenue at $2.373b, or 62% of total revenue (p. F-31). It reports 2015 total revenue at $1.995b (p. 96), so we estimate $1.273b ridesharing revenue in US and Canada, using the 62% figure from 2016. 2016 ridesharing revenue as a percentage of bookings was 18% (p. 122), yielding gross US and Canada bookings of $13.2b and $7.1b in 2016 and 2015 respectively.

Thus Lyft had approximately 6% of Uber's bookings in 2015 ($0.4b/$7.1b) and 14% of Uber's bookings in 2016 ($1.9b/$13.2b).

A1.2 Uber US market capitalization and gross bookings

The spreadsheet printed in Figure A1 details our calculations of Uber's 2019 US ridesharing-based market capitalization and 2019 US ridesharing bookings (i.e. total gross revenues collected from riders).

Figure A1: Calculations for Uber US market capitalization and bookings

	A	B	C	D	E
1			UBER US market capitalization calculation		
2		**Shares**	**Average share price**	**Market cap**	**Notes**
3	**Q3 2019**	1,700,213,000	$30.47	$51,805,490,110	PRODUCT(B3, C3)
4	**Q4 2019**	1,710,260,000	$29.74	$50,863,132,400	PRODUCT(B4, C4)
5	**2019 average**			$51,334,311,255	AVERAGE(D3, D4)
6					
7	**US/Canada share of revenue**			62%	
8	**Ridesharing share of revenue**			75%	
9	**US share of US/Canada**			90%	
10	**Non-NYC/SEA share**			89%	
11					
12	**Average market cap (US)**			$19,164,843,965	PRODUCT(D5:D10)
13					
14			UBER US gross bookings calculation		
15			**Assumed average ride price**	**Bookings (2019)**	
16	**Total ridesharing bookings**			$49,700,000,000	
17	**US ridesharing bookings**			$27,732,600,000	PRODUCT(D7, D9, D16)
18	**NYC bookings (circa 2019)**		$15.00	$2,737,500,000	500,000 daily rides
19	**SEA bookings (circa 2019)**		$10.00	$255,500,000	70,000 daily rides
20	**US bookings net of NYC/SEA**			$24,739,600,000	D17 - (D18 + D19)
21					
22					
23	**Sources:**	UBER quarterly earnings reports	https://www.wsj.com/market-data/quotes/UBER/advanced-chart	UBER 2019 annual report (pp. 61, 115); NYC and Seattle documents	

A2 Consumer surplus formula

We consider the following demand curve featuring a constant elasticity of demand β_1, with $\beta_1 < -1$ (i.e. demand is elastic):

$$ln(q) = \beta_0 + \beta_1 ln(p)$$

Then:

$$\begin{aligned}
ln(p) &= -\frac{\beta_0}{\beta_1} + \frac{1}{\beta_1} ln(q) \\
p &= e^{-\frac{\beta_0}{\beta_1} + \frac{1}{\beta_1} ln(q)} \\
&= e^{-\frac{\beta_0}{\beta_1}} q^{\beta_1^{-1}} \\
&= a q^{\beta_1^{-1}}
\end{aligned}$$

where $a = e^{-\frac{\beta_0}{\beta_1}}$. Total area under the demand curve at ridership level $\bar{q}$ is:

$$\begin{aligned}
\int_0^{\bar{q}} a q^{\beta_1^{-1}} dq &= \frac{a q^{\beta_1^{-1}+1}}{\beta_1^{-1}+1} \Big|_0^{\bar{q}} \\
&= \frac{a \bar{q}^{\beta_1^{-1}+1}}{\beta_1^{-1}+1}
\end{aligned}$$

Consumer surplus relative to the counterfactual of no Uber service is then:

$$\begin{aligned}
\frac{a \bar{q}^{\beta_1^{-1}+1}}{\beta_1^{-1}+1} - \bar{p}\bar{q} &= \frac{a \bar{q}^{\beta_1^{-1}+1}}{\beta_1^{-1}+1} - a \bar{q}^{\beta_1^{-1}+1} \\
&= a \bar{q}^{\beta_1^{-1}+1} \left(\frac{1}{\beta_1^{-1}+1} - 1\right) \\
&= -\frac{1}{1+\beta_1} \bar{p}\bar{q}
\end{aligned}$$

Thus consumer surplus equals total revenue (collected from riders) rescaled by $-\frac{1}{1+\beta_1}$.

Using a constant elasticity of demand down to a quantity of zero implies, however, that there is no choke price at which demand falls to zero (indeed, this formula fails when demand becomes inelastic). To prevent willingness to pay from becoming unbounded, we impose a choke price of $5\bar{p}$, similar to Cohen et al. (2016). At a choke price $5\bar{p}$, we net out the consumer surplus corresponding to revenues that would be collected if the price were $5\bar{p}$ (as there is no additional consumer surplus beyond that price). Revenues that would be collected at the choke price of $5\bar{p}$ are:

$$5\bar{p} \cdot q(5\bar{p}) = 5\bar{p} \cdot 5^{\beta_1}\bar{q} = 5^{\beta_1+1}\bar{p}\bar{q}$$

Thus, total consumer surplus equals total revenue rescaled by $-\frac{1}{1+\beta_1}(1-5^{\beta_1+1})$.

For example, if the price elasticity of demand is $\beta_1 = -8$, then consumer surplus is effectively one-seventh of total revenue, as $-\frac{1}{1-8} = \frac{1}{7}$ and $(1-5^{-8+1}) \approx 1$. If $\beta_1 = -2$, then consumer surplus is 80% of total revenue, as $-\frac{1}{1-2} = 1$ and $(1-5^{-2+1}) = 0.8$.

Figure A2: US Traffic Fatalities by Hour of Day, 2012-16

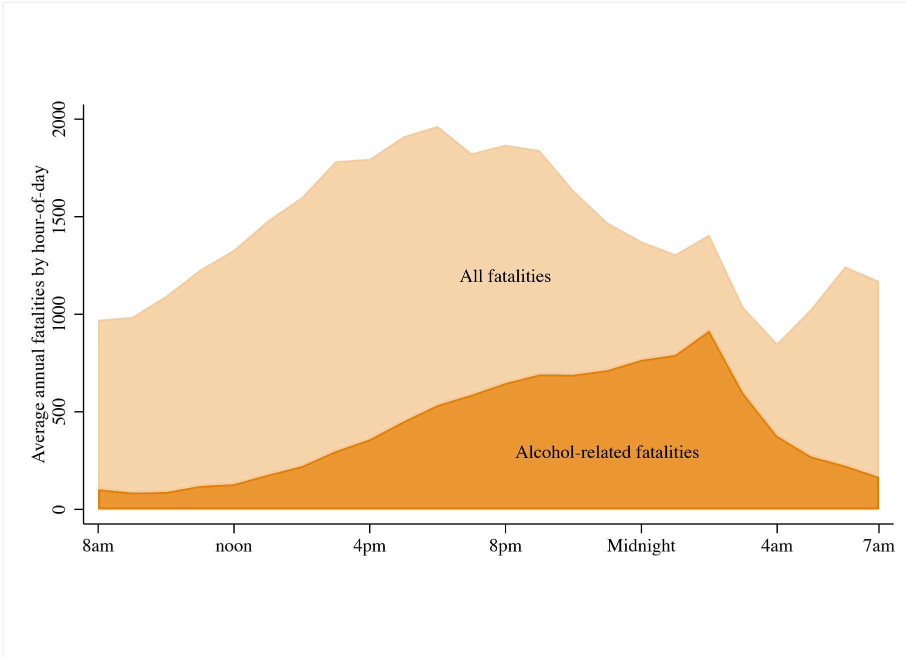

Figure A3: US Traffic Fatalities in Cities by Hour of Day, 2012-16

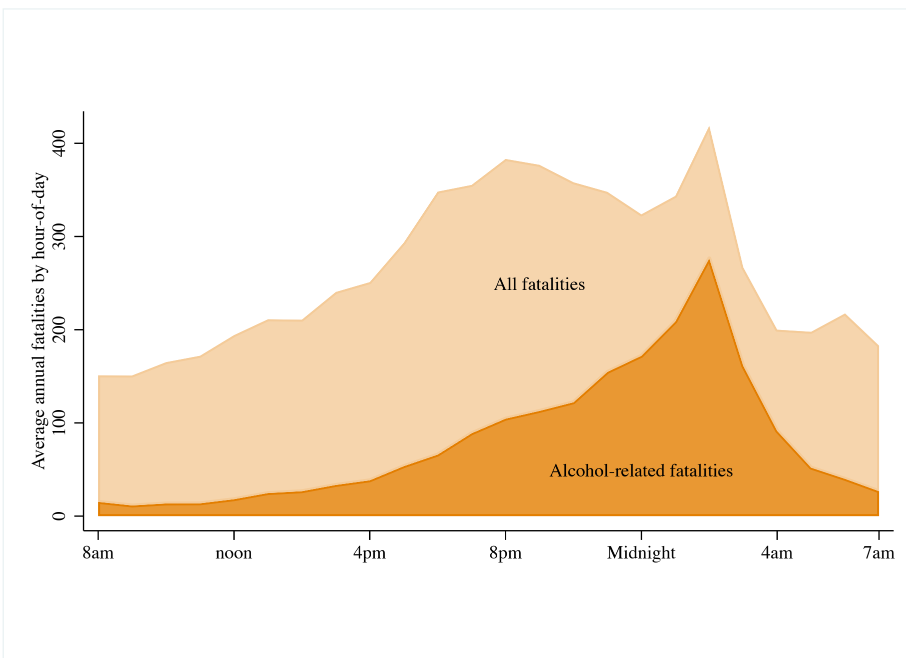

Table A1: Summary Statistics

	(1) All tracts	(2) Uber tracts
Census tracts	73,057	44,808
Tract-month observations	4,383,420	2,688,480
A. All fatal accidents		
Annual fatalities	34,077	15,898.4
Probability of fatal crash (tract-month)	0.034	0.027
Annual single-vehicle fatalities	14,002.2	5,302.4
Annual multi-vehicle fatalities	14,003.2	6,422.8
Annual pedestrian/cyclist fatalities	6,071.6	4,173.2
B. Alcohol-involved fatal accidents		
Annual fatalities	10,035.4	4726.2
Probability of fatal crash (tract-month)	0.010	0.008
Annual single-vehicle fatalities	6,016	2,463.8
Annual multi-vehicle fatalities	3,500.4	1,902.2
Annual pedestrian/cyclist fatalities	519	360.2

Notes: This table reports summary statistics from our FARS dataset. The unit of observation is the Census tract by month. The time period spans January 2012 to December 2016. Column (1) reports statistics for all Census tracts, and Column (2) reports statistics for Census tracts that see any Uber activity by December 2016.

Table A2: The Effect of Uber on Alcohol-Related Traffic Fatalities, Distance$^{-0.9}$ Weight

	(1)	(2)	(3)	(4)	(5)
	Drunk death	Drunk death	Drunk death	Drunk death	Drunk death
Rideshare index	-0.041	-0.047	-0.044	-0.044	-0.035
	(0.013)	(0.013)	(0.014)	(0.015)	(0.013)
	0.007	*0.003*	*0.006*	*0.008*	*0.014*
Observations	2,688,480	2,206,800	1,098,840	657,900	438,840
Max tract rideshare activity	Nonzero	Top 50 pct	Top 25 pct	Top 15 pct	Top 10 pct
Tracts	44,808	36,780	18,314	10,965	7,314
Mean dependent variable	0.788	0.705	0.609	0.556	0.497
Mean index	0.119	0.145	0.279	0.435	0.600

Notes: This table reports coefficient estimates from five separate least squares regressions that progressively restrict the sample to locations with higher ridesharing activity by the end of our sample period. The unit of observation is the Census tract by month. The dependent variable in all regressions is an indicator for alcohol-related fatalities, multiplied by 100. The independent variable is the weighted average of rideshare activity originating within a 10-mile radius of Census tract i in month t, with weights equal to $0.25 \cdot \text{distance}^{-0.9}$, normalized such that a value of 1 corresponds to the average value of the index in 2019. All regressions include Census-tract and month-of-sample fixed effects. Parentheses contain standard errors clustered by state. FDR-control q-values in *italics*.

Table A3: The Effect of Uber on Alcohol-Related Traffic Fatalities, Distance$^{-1.1}$ Weight

	(1) Drunk death	(2) Drunk death	(3) Drunk death	(4) Drunk death	(5) Drunk death
Rideshare index	-0.035	-0.039	-0.036	-0.035	-0.026
	(0.011)	(0.011)	(0.011)	(0.012)	(0.010)
	0.007	*0.003*	*0.006*	*0.011*	*0.021*
Observations	2,688,480	2,206,800	1,098,840	657,900	438,840
Max tract rideshare activity	Nonzero	Top 50 pct	Top 25 pct	Top 15 pct	Top 10 pct
Tracts	44,808	36,780	18,314	10,965	7,314
Mean dependent variable	0.788	0.705	0.609	0.556	0.497
Mean index	0.120	0.146	0.281	0.439	0.610

Notes: This table reports coefficient estimates from five separate least squares regressions that progressively restrict the sample to locations with higher ridesharing activity by the end of our sample period. The unit of observation is the Census tract by month. The dependent variable in all regressions is an indicator for alcohol-related fatalities, multiplied by 100. The independent variable is the weighted average of rideshare activity originating within a 10-mile radius of Census tract i in month t, with weights equal to $0.25 \cdot \text{distance}^{-1.1}$, normalized such that a value of 1 corresponds to the average value of the index in 2019. All regressions include Census-tract and month-of-sample fixed effects. Parentheses contain standard errors clustered by state. FDR-control q-values in *italics*.

Table A4: The Effect of Uber on Alcohol-Related Traffic Fatalities, State-by-Month FEs

	(1) Drunk death	(2) Drunk death	(3) Drunk death
Rideshare index	-0.049	-0.032	-0.039
	(0.011)	(0.012)	(0.011)
	0.001	*0.016*	*0.003*
Observations	2,206,800	2,206,800	1,098,780
Max tract rideshare activity	Top 50 pct	Top 50 pct	Top 25 pct
Dependent variable coverage	All	8p-6a	All
Tracts	36,780	36,780	18,314
Mean dependent variable	0.705	0.506	0.609
Mean index	0.145	0.145	0.280

Notes: This table reports coefficient estimates from three separate least squares regressions. Specifications are identical to Tables 2 and 3 except that all regressions include state-by-month-of-sample fixed effects. Parentheses contain standard errors clustered by state. FDR-control q-values in *italics*.

Table A5: The Effect of Uber on Alcohol-Related Traffic Fatalities, time-shifted placebos

	(1) Drunk death	(2) Drunk death	(3) Drunk death	(4) Drunk death
Rideshare index shifted 7/12 yrs	0.009	0.025	-0.024	-0.019
	(0.020)	(0.017)	(0.038)	(0.043)
	0.787	*0.343*	*0.787*	*0.787*
Observations	2,206,800	1,098,840	1,765,440	879,072
Max tract rideshare activity	Top 50 pct	Top 25 pct	Top 50 pct	Top 25 pct
Sample years	2005-09	2005-09	2001-04	2001-04
Tracts	36,780	18,314	36,780	18,314
Mean dependent variable	0.938	0.852	0.960	0.922
Mean index	0.145	0.280	0.181	0.349

Notes: This table reports coefficient estimates from four separate least squares regressions. The unit of observation is the Census tract by month, and the estimation sample covers January 2001 to December 2005 or January 2006 to December 2010. The dependent variable in the all columns is an indicator for alcohol-related fatalities, multiplied by 100. The independent variable is the weighted average of rideshare activity originating within a 10-mile radius of Census tract i in month t, with weights equal to $0.25 \cdot \text{distance}^{-1}$, shifted back in time by 7 or 12 years. All regressions include Census-tract and month-of-sample fixed effects. Parentheses contain standard errors clustered by state. FDR-control q-values in *italics*.

Table A6: The Effect of Uber on All Traffic Fatalities, time-shifted placebos

	(1)	(2)	(3)	(4)
	Any death	Any death	Any death	Any death
Rideshare index shifted 7/12 yrs	0.053	0.056	-0.057	-0.020
	(0.031)	(0.027)	(0.054)	(0.058)
	0.383	*0.343*	*0.708*	*0.793*
Observations	2,206,800	1,098,840	1,765,440	879,072
Max tract rideshare activity	Top 50 pct	Top 25 pct	Top 50 pct	Top 25 pct
Sample years	2005-09	2005-09	2001-04	2001-04
Tracts	36,780	18,314	36,780	18,314
Mean dependent variable	2.582	2.372	2.554	2.394
Mean index	0.145	0.280	0.181	0.349

Notes: This table reports coefficient estimates from four separate least squares regressions. The unit of observation is the Census tract by month, and the estimation sample covers January 2001 to December 2005 or January 2006 to December 2010. The dependent variable in all columns is an indicator for any fatalities, multiplied by 100. The independent variable is the weighted average of rideshare activity originating within a 10-mile radius of Census tract i in month t, with weights equal to $0.25 \cdot \text{distance}^{-1}$, shifted back in time by 7 or 12 years. All regressions include Census-tract and month-of-sample fixed effects. Parentheses contain standard errors clustered by state. FDR-control q-values in *italics*.

Table A7: The Effect of Uber on Alcohol-Related Traffic Fatalities by crash type

	(1) 1-vehicle drunk death	(2) 2-vehicles drunk death	(3) Pedestrian drunk death
Rideshare index	-0.013	-0.023	-0.007
	(0.006)	(0.005)	(0.006)
	0.029	*0.001*	*0.250*
Observations	2,206,800	2,206,800	2,206,800
Tracts	36,780	36,780	36,780
Mean dependent variable	0.360	0.283	0.0660
Mean index	0.145	0.145	0.145

Notes: This table reports coefficient estimates from three separate least squares regressions. The unit of observation is the Census tract by month. The dependent variable in all regressions is an indicator for alcohol-related fatalities (by crash type), multiplied by 100. The independent variable is the weighted average of rideshare activity originating within a 10-mile radius of Census tract i in month t, with weights equal to $0.25 \cdot \text{distance}^{-1}$. All regressions include Census-tract and month-of-sample fixed effects. Parentheses contain standard errors clustered by state. FDR-control q-values in *italics*.