

UC Santa Barbara

Econ 196 Honors Thesis

Title

Driving You to Drink: Do Ride-Hailing Applications Affect Restaurant and Bar Revenue?

Permalink

<https://escholarship.org/uc/item/0322526g>

Author

Plesset, Aaron

Publication Date

2018-06-26

Data Availability

The data associated with this publication are within the manuscript.

Undergraduate

Driving You to Drink: Do Ride-Hailing Applications Affect Restaurant and Bar Revenue?

Advisor: Doug Steigerwald

March 18, 2018

Abstract

There is recent evidence that Ride-Hailing applications reduce the number of alcohol-related motor vehicle collisions and fatalities. This paper aims to investigate how they may impact the restaurant and bar industry by estimating a difference-in-differences model across cities in California. Results suggest an average increase of \$36 in quarterly per capita restaurant and bar revenue. This result is robust to model validation where the three largest cities are separately excluded from the model. Additional robustness checks do not provide any evidence that the parallel trends assumption is violated.

1 Introduction

Ride-hailing applications such as UberX and Lyft have been a hot topic of discussion since their introduction in 2012. Given the disruption that ride-sharing has caused in the taxi industry, policy makers have been prompted to consider regulations that could mitigate the upheaval that it has caused. For example, the European Court of Justice recently ruled that Uber should be regulated as a transportation service rather than a platform that connects drivers to people who need a ride¹. This provides additional hurdles for potential drivers. Additionally, it creates a cap on the number of available Uber drivers on the road because some cities issue a limited number of permits. Pressure to regulate has not been without resistance however, given that ride-sharing may provide economic and social benefits such as increased market efficiency² and a decrease in the number of alcohol-related driving incidences. For example, Morrisson et al. (2017) uses the fact that Uber services were paused and resumed in four major U.S. cities to see how this affects the number of alcohol-related motor vehicle collisions. In some cases, the alcohol-related collision rate decreased by as much as 61.8%. Perhaps more importantly, Greenwood (2017) uses difference-in-differences methodology across the State of California to see if the introduction of Uber is associated with a decrease in the number of alcohol-related motor vehicle fatalities. It finds that on average, Uber's introduction may contribute to a 3.6% reduction in the number of alcohol-related driving fatalities. Naturally, it would be interesting to extend these results on alcohol-related behaviors by exploring whether ride-hailing applications have also impacted alcohol consumption.

Although it would be interesting to study the impact of ride-hailing applications on individual alcohol consumption patterns, this paper focuses mainly on its potential effect on restaurant and bar revenue. In California, where taxis are relatively scarce, ride-hailing applications have solved a number of financial and logistical issues associated with a night of drinking. Previously, the difficulty of ride coordination may have been a limiting factor for groups of people who wanted to go out. Finding a ride to and from a venue, whether it be relying on a friend or hailing a cab, is something that is inherently not an easy task. Furthermore, for those who were successful at coordinating a ride, the extent to which bar-hopping took place may have been limited by proximity between bars. In addition to lowering monetary costs associated with a night of drinking, ride-hailing applications have effectively eliminated non-monetary costs and even bared efficiency gains over traditional taxis³. Thus, we may expect to see increases in restaurant and bar revenue resulting from their introduction.

Because governments receive a share of business revenue, changes in restaurant and bar revenue resulting from ride-hailing applications may be of interest to policy makers and business owners alike. On one hand, additional revenue is clearly beneficial to cities. On the other hand, revenue generated from increased alcohol consumption may entail negative health-related externalities. As previously mentioned, this paper mainly focuses on aspects related to changes in revenue. In my analysis, I employ similar methods to those utilized by the existing literature in an attempt to assess whether ride-hailing applications impact restaurant and bar revenue. Specifically, I utilize difference-in-differences methodology across 216 cities in California to determine if there is evidence of a relationship between the introduction of Uber and changes in restaurant and bar revenue per capita. Results suggest that on average, per capita restaurant and bar revenue increases by approximately \$36 per quarter following the introduction of Uber. This result

¹<http://www.latimes.com/business/technology/la-fi-tn-uber-european-union-ruling-20171220-story.html>

²Cramer et al. (2016)

³Cramer et al. (2016)

is robust to model validation where the three largest cities in California are separately excluded from the analysis.

2 Methodology

An ideal experiment to measure the effect ride-hailing applications have on alcohol consumption would be a scenario similar to what follows. Prior to ride-hailing applications, suppose we had two cities with the same characteristics that determine alcohol consumption patterns. For these two cities, we observe that on average the difference in alcohol consumption per capita between them is some constant K . So, we randomly introduce a ride-hailing service to exactly one of the two cities at time t . If at time $t + i$, where $i > 0$, we observe that the difference in alcohol consumption per capita between the city that received Uber and the city that did not is no longer K but instead $K + \delta$, we can attribute this difference in differences δ to the introduction of Uber (assuming that nothing else has changed).

The ideal real world experiment is generally unattainable, however many researchers in social sciences mimic it through empirical techniques that allow us to make the same causal inference. This methodology is referred to as difference-in-differences. Common applications include studying the effect of introducing a new policy, service, or event of this nature on some response variable. An example of difference-in-differences is Card and Krueger (1994) who study the effect of a minimum wage change in New Jersey on employment. Although we generally do not have two cities with the same event-related characteristics, we can control for differences in these characteristics through regression analysis. For example, it could be the case that ride-hailing applications are more likely to influence alcohol consumption patterns for a city where the median age is 25 versus a city where the median age is 40. Thus, this is something I control for.

Because we cannot control for every characteristic, for our estimates to be unbiased the remaining differences between each group must be constant constant over time. Furthermore, the impact of time must be uniform across all groups. For example, let's say that during the treatment period, a non-trivial amount of cities introduce an alcohol tax increase whereas the remaining do not. If the majority of these cities lie in the treatment group, we may expect our estimate of the treatment effect to be too low (assuming people consume less alcohol when prices increase). Parameters for these latent variables are included in my model. Lastly, an obvious bias to the estimation could result from the difference between treatment and control not being a constant K but instead one that changes over time. The assumption that the difference is constant, the parallel trends assumption, cannot be directly tested. I do however look for evidence that it is violated through Granger causality testing discussed in section 2.1.

2.1 Model

To determine the effect Uber has on a city's level of alcohol consumption, I estimate the following difference-in-differences model:

$$revenue_{it} = \alpha_i + \lambda_t + \delta treated_{it} + \gamma_1 median.age_{it} + \gamma_2 median.income_{it} + \epsilon_{it} \quad (1)$$

Where :

- Revenue is quarterly per capita restaurant and bar revenue for city i at time t
- α_i represents the city fixed effects for city i

- λ_t represents the time fixed effects for time t
- Treated is an indicator variable equal to 1 if city i is treated at time t , 0 otherwise
- Median Age is the median age for the county in which city i resides
- Median Income is the median income for the county in which city i resides

**For notation, I will denote $\mathbf{X} = \langle \text{median.age}_{it}, \text{median.income}_{it} \rangle$ as the vector of the control covariates. Similarly, let $\gamma = \langle \gamma_1, \gamma_2 \rangle$*

So that

$$\mathbb{E}[\text{revenue}_{it} | \text{city}_i, \text{time}_t, \mathbf{X}, \text{treated} = 1] - \mathbb{E}[\text{revenue}_{it} | \text{city}_i, \text{time}_t, \mathbf{X}, \text{treated} = 0] = \delta$$

*Note: City fixed effects estimate the constant city differences over time. Similarly, time fixed effects estimate the uniform impact time has on sales revenue across all cities.

Since Uber service was not randomly allocated to different cities at different times, there may be endogeneity concerns. Factors that may endogenize the introduction of Uber are likely associated with the population, age, and income distribution of a city and are thus controlled for. Nonetheless, I address any anticipatory effect concerns using Granger causality testing. To do this, I estimate a model similar to that used in Autor (2003):

$$\text{revenue}_{it} = \alpha_i + \lambda_t + \sum_{\tau=0}^5 \beta_{-\tau} \text{introduction}_{i(t-\tau)} + \sum_{\tau=1}^2 \beta_{\tau} \text{introduction}_{i(t+\tau)} + \gamma \mathbf{X} + \epsilon_{it} \quad (2)$$

In this model, introduction is an indicator variable equal to one if period $t + \tau$ is the period in which Uber is first introduced for city i . The remaining variables have the same representation that they have in model (1). Hence, we see that for $\tau < 0$, β_{τ} represents the effect of the treatment $|\tau|$ period(s) after the treatment has begun. For $\tau > 0$, β_{τ} represents the “treatment effect” τ period(s) prior to being treated⁴. Clearly, $\beta_{|\tau|}$ being non-zero would indicate that it was not the treatment per se that caused changes in alcohol sales per capita since the treatment has not yet started.

There are two ways in which $\beta_{|\tau|}$ might be non-zero. First, it’s definitely plausible that ride-hailing companies would select cities where alcohol consumption was higher since this may generate more revenue. This would indicate that any effect I found would be correlative and not causal. Second, if alcohol revenue were to be systematically increasing more in treatment groups relative to control groups, this result would surface by $\beta_{|\tau|}$ being non-zero. In other words, if the parallel trends assumption is violated, we would expect the treatment group to deviate from the counterfactual prior to the treatment beginning.

Five lags were chosen for this model because that is the average length of the treatment window. Two leads were included because the estimates are means of robustness check and are not important in and of themselves.

⁴Autor (2003) refers to these as lags and leads, respectively

Table 1:

Arrival		Periods	
Region	Date	Region	Date
Bay Area	2012 Q3	Los Angeles	2013 Q3
Orange County	2013 Q3	San Bernardino	2014 Q1
San Diego	2013 Q3	Riverside	2014 Q1
Sacramento	2013 Q4	Fresno	2014 Q2
Santa Barbara	2014 Q1	Stanislaus	2014 Q3

Lastly, the initial analysis included all cities. Upon inspection of the data, the city of Industry was highly unrepresentative of the treatment group and was a source of significant bias in the estimation. Mainly, it has exceptionally few people who actually reside in the city causing revenue per capita to be seventy times greater than the next highest sales per capita for any city. Since Industry was in the treatment group, its inclusion resulted in a ten-fold increase in the estimate for the treatment effect.

3 Data

3.1 Dependent Variable

Sales revenue from the “Eating and Drinking Places” sector can be found on the California Board of Equalization website. Although this sector includes food services in addition to drinking places, there should be no reason that the arrival of ride-hailing impacts food consumption patterns when this is not accompanied by alcohol. Thus, the effect I find from taking differences should mainly reflect differences in alcohol consumption that may or may not be accompanied by food. Since I am interested in alcohol consumption per capita, I divide the sales revenue data by corresponding quarterly population data obtained from the California Department of Finance. Both data sets are at the city level broken down by quarter.

3.2 Uber’s Arrival

I compiled Uber’s arrival date for all of the treated counties using several sources. First, I checked Uber’s blog to see if they made an announcement that they would be arriving in a particular area. This was the case for several counties. Next, I used an online source⁵ that records snapshots of websites over time. I used snapshots of the Uber website to pinpoint the exact quarter that Uber arrived for a majority of the counties. Assuming Uber posted a region to its website within a month of its availability in that location, there should be no measurement error. Records were sparse for a couple regions on the web archive so the last sources that I used were news articles that I found online.

3.3 Covariates

In the analysis, I control for the median age and median income for each city. Median age data at the county level come from the U.S. Census Bureau website (ACS Survey), and median income data at the county level come from the California Franchise Tax Board website. Median age data for cities Hollister and Red Bluff, California were not available so these cities were not included in the analysis.

Table 2:

Statistic	N	Mean	St. Dev.	Min	Max
Population	8,640	111,951	285,141	8,377	3,945,037
Median Income	8,640	35,891	6,813	22,841	59,896
Median Age	8,640	35	2.886	29	50
Revenue	8,640	490	344.144	22	2,995

Table 2 shows summary statistics for the covariates. The intersection of these datasets spans 2005 - 2014

4 Results

4.1 Empirical Findings

The main finding is that the arrival of Uber is on average associated with a \$36 dollar per quarter increase in restaurant and bar revenue per capita. Model validation in which the three largest cities in California⁶ are non-jointly excluded from the estimation is used to show that this result is not single-handedly being driven by a major California city. Furthermore, the number of housing starts in California has been rising since the financial crisis of 2008⁷. Since this phenomenon may be disproportionately higher in highly populated cities, it could potentially be confounding our estimate. The most obvious aspect of more housing starts that could drive alcohol sales is the fact that population is growing faster, however the per capita specification of the regression model accounts for this. Nonetheless, housing starts are a widely used indicator of economic conditions. Disproportionate economic conditions may cast some doubt on the assumption that time impacts cities uniformly. Excluding large California cities allows us to see how variable our estimate is to the exclusion of cities where we might expect to see many housing starts.

⁵<http://web.archive.org>

⁶Los Angeles, San Diego, San Jose

⁷http://archive.senate.ca.gov/sites/archive.senate.ca.gov/files/committees/2015-16/stran.senate.ca.gov/sites/stran.senate.ca.gov/files/ppic_report_jan_2016.pdf

Table 3:

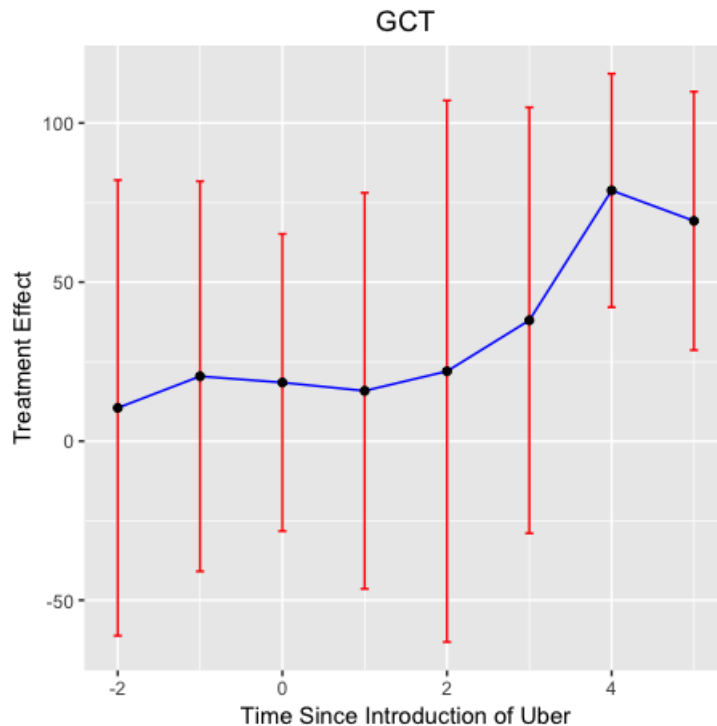
	<i>Dependent variable:</i>			
	All Cities	Revenue		
		-San Diego	-Los Angeles	-San Jose
	(1)	(2)	(3)	(4)
Median Income	-0.003*** (0.0004)	-0.003*** (0.0004)	-0.003*** (0.0004)	-0.003*** (0.0004)
Median Age	29.648*** (0.638)	29.720*** (0.641)	29.764*** (0.643)	29.185*** (0.644)
Treatment	35.808*** (9.718)	35.040*** (9.741)	36.386*** (9.825)	41.520*** (10.162)
Observations	8,640	8,600	8,600	8,600
R ²	0.042	0.043	0.043	0.043
Adjusted R ²	0.038	0.038	0.038	0.038

Note:

*p<0.1; **p<0.05; ***p<0.01

Table 3 shows estimates for the four regressions. Column (1) estimates the difference-in-differences model for all cities, columns (2)-(4) show the effect on the estimates from excluding the city corresponding to the column name. All standard errors reported are heteroskedasticity robust and clustered at the city level.

Looking at the regression table, the direction for Median Income and Median Age's effect on restaurant and bar revenue is surprising. One would expect that cities with wealthier individuals are able to spend more money at restaurants and bars. Similarly, we may expect that younger individuals generate more restaurant and bar revenue than older individuals. That being said, the obtained results are not implausible. The direction for Median Income may be justified by the fact that the variable measures the income of individuals who reside in a particular city, not individuals who partake in its nightlife. Regarding the direction for Median Age, two stories come to mind. First, imagine that a city has many families. Because children are included in the census, we may expect these cities to have a lower Median Age relative to a city with fewer families. The city with many families may also have lower restaurant and bar revenue because parents tend to go out less frequently and because there are proportionally fewer people relative to the city's population who are of drinking age. The other story is that older people enjoy more leisure time than younger people. Since restaurant revenue is included in our revenue measure, we may be measuring this to some extent. Of course, the two scenarios are not mutually exclusive so a combination of them may also be possible. Further investigation will need to be done to more accurately explain this phenomenon.



Lastly, the results of the Granger Causality testing do not indicate any anticipatory effects for the treatment. In other words, zero is well within the 95% confidence interval for the estimates of the leads. Additionally, Figure 1 reveals that the effect from the introduction of Uber seems to take off after the second lag, however the effect does not become statistically significant until sometime after the third lag (approximately nine months after the initial treatment). This is not particularly surprising, seeing as it takes time for a network of drivers to become established. Moreover, UberX was still a relatively new service during the treatment window. Because people need to know about a service in order to use it and know what it is useful for, this could have been a source of delay for the treatment effect. Estimates for model (2) can be found in Table 4.

4.2 Discussion

From a revenue perspective, the main finding is promising. Ride-hailing applications appear to have successfully generated additional revenue in the Food Services and Drinking places sectors. Because taxi services and ride-hailing applications superficially seem to be perfect substitutes, one may wonder whether these results extend to cities such as New York City where taxi services are ubiquitous. If ride-hailing applications and taxi services were truly perfect substitutes however, we might not expect to see ride-hailing applications cause much behavioral change in New York City where taxi services are already very common. Contrary to this idea however, Peck (2017) studies the impact of the introduction of Uber across New York City's boroughs on alcohol-related motor vehicle collisions using difference-in-differences and synthetic control methodologies. She finds an average decrease of 40 alcohol-related collisions per month. This is consistent with previous literature where analysis was performed across the state of California (Green 2017) and across major cities in multiple states (Morisson et al. 2017).

Clearly, an increase in Restaurant and Bar Revenue may have implications that go beyond benefits for city revenue and that for certain types of business. For example, there are health consequences associated with an increase in alcohol consumption. In extreme circumstances where the result was entirely driven by changes in alcohol consumption, we would observe that on average people are spending approximately \$12 more per month. Based on current prices, this may amount to an extra two drinks per month for an individual. That being said, it is unlikely that this effect is entirely driven by changes in alcohol consumption since people do not typically consume alcohol in isolation and since the revenue is aggregated across both the Food Services and Drinking Places sectors. The former contains locations that certainly serve prepared food and the latter contains some that do and some that do not. As previously mentioned, this paper does not attempt to assess health-consequences related to the introduction of ride-hailing. This may however be an interesting avenue for future research.

5 Conclusion

Uber services were introduced at different times in different cities. Under some assumptions, these events emulate a real-world experiment where we randomly decide to introduce Uber services to similar cities. Accordingly, this paper estimates a difference-in-differences model to determine the average effect that Uber - and possibly other ride-hailing applications - have had on the restaurant and bar industry for each city. The assumptions of the model are that time impacts each city uniformly, there is a constant difference between each city that is time invariant, and lastly that difference in rate of change over time for revenue between treated and untreated cities is constant prior to the treatment. Fixed effects allow us to estimate the constant differences between cities and estimate the effect time has on cities. We use the last assumption to determine a counter-factual for where revenue would otherwise be had a city not received Uber services.

In estimating the model, I find that the introduction of Uber is on average associated with a \$36 increase in restaurant and bar revenue per quarter. This result varies little when major cities are separately excluded from the model. Robustness checks were performed in search of evidence that the parallel trends assumption was violated, something that could potentially invalidate the findings. In essence, the robustness check was an estimation of a model that tries to capture the effect of Uber arriving in the near future. This effect is estimated by the leads in model (2). If it were the case that the difference in rate of change for revenue between treated and untreated cities was not constant prior to the treatment, we would expect to see this surface in the model estimation since the treated cities would effectively be leaving the counterfactual before the treatment has actually started. Zero was well within a 95% confidence interval for the estimation of the leads indicating that there is no evidence for a violation of the parallel trends assumption. When estimating this model, lags were also included which effectively estimates the effect that Uber has on revenue given that it arrived recently. Specifically, indicator variables for distance from the treatment introduction date give insight into the effect Uber has on restaurant and bar revenue over time. Uber's effect on quarterly restaurant and bar revenue per capita does not appear to be immediate, and begins to stand out approximately nine months after Uber's arrival.

Finally, previous literature suggests that there are fewer incidences of drunk-driving that result from Uber's introduction. In conjunction with that, we may be able to infer from our result that people might be drinking more because they do not have to

drive. Naturally, this may seem like a negative consequence ride-hailing applications entail regarding policy implications. That being said, the magnitude of the result may not be cause of alarm because it only amounts to an average per capita increase of \$12 per month spent at restaurant and bars. Furthermore, since alcohol is not always consumed in isolation, this result is not entirely driven from increases in alcohol consumption. Further research would need to be done to determine the extent to which this result is driven solely from drinking. All in all however, the magnitude of the per-capita increase in restaurant and bar revenue may be seen as a positive effect that ride-hailing applications have had economically.

References

- Wesley A. Austin and Rand W. Ressler. Do designated drivers and workplace policies effect alcohol consumption? *Journal of Behavioral and Experimental Economics (formerly The Journal of Socio-Economics)*, 41(1):104–109, 2012. 5
- David H. Autor, Frank Levy, and Richard J. Murnane. The skill content of recent technological change: An empirical exploration*. *The Quarterly Journal of Economics*, 118(4):1279–1333, 2003. 5
- David Card, Lawrence F. Katz, and Alan B. Krueger. Comment on david neumark and william wascher, employment effects of minimum and subminimum wages: Panel data on state minimum wage laws. *ILR Review*, 47(3):487–497, 1994. 5
- Judd Cramer and Alan B. Krueger. Disruptive change in the taxi business: The case of uber. *American Economic Review*, 106(5):177–82, May 2016. 5
- Brad N Greenwood. Show me the way to go home an empirical investigation of ride-sharing and alcohol related motor vehicle fatalities. *Management information systems : mis quarterly*, 41(1):163–187, 2017. ISSN 02767783. 5
- Christopher N Morrison, Sara F Jacoby, Beidi Dong, M Kit Delgado, and Douglas J Wiebe. Ridesharing and motor vehicle crashes in four u.s. cities: An interrupted time series analysis. *American journal of epidemiology*, June 2017. ISSN 1476-6256. 5
- Jessica Lynn Peck. *New York City Drunk Driving After Uber (Working Paper No. 13)*. CUNY, 2017. 5
- Cramer and Krueger [2016], Greenwood [2017], Morrison et al. [2017], Peck [2017], Austin and Ressler [2012], Autor et al. [2003], Card et al. [1994]