A Quasi-experimental Estimate of the Impact of P2P Transportation Platforms on Urban Consumer Patterns

Zhe Zhang
Carnegie Mellon University
5000 Forbes Ave
Pittsburgh, PA 15213
zhezhang@cmu.edu

Beibei Li
Carnegie Mellon University
5000 Forbes Ave
Pittsburgh, PA 15213
beibeili@andrew.cmu.edu

ABSTRACT

With the pervasiveness of mobile technology and location-based computing, new forms of smart urban transportation, such as Uber & Lyft, have become increasingly popular. These new forms of urban infrastructure can influence individuals' movement frictions and patterns, in turn influencing local consumption patterns and the economic performance of local businesses. To gain insights about future impact of urban transportation changes, in this paper, we utilize a novel dataset and econometric analysis methods to present a quasi-experimental examination of how the emerging growth of peer-to-peer car sharing services may have affected local consumer mobility and consumption patterns.

CCS CONCEPTS

• **Applied computing** → **Economics**; Transportation; • **Information systems** → *Location based services*;

ACM Reference format:
Zhe Zhang and Beibei Li. 2017. A Quasi-experimental Estimate of the Impact of P2P Transportation Platforms on Urban Consumer Patterns. In *Proceedings of KDD '17, August 13-17, 2017, Halifax, NS, Canada, ,* 9 pages.
DOI: 10.1145/3097983.3098058

1 INTRODUCTION

With the pervasiveness of mobile technology and location-based computing, new forms of smart urban transportation, such as Uber & Lyft — peer-to-peer car transportation services — have become increasingly popular. These new forms of urban infrastructure and transportation can influence individuals' movement frictions, in turn influencing local consumption patterns and the economic performance of local businesses. In this paper, we utilize a novel dataset to present a quasi-experimental examination of how the emerging growth of peer-to-peer car sharing services may have affected local consumer mobility and consumption patterns (e.g. frequency of consumption, total amount spent, and diversity of consumption).

Under KDD's Discovery theme, as cities adapt to continuing changes in transportation technology, it is of interest to study how such changes to the transportation system may affect consumers, local businesses, and social mobility patterns. This research can provide social insights for planners and government officials considering future changes, such as self-driving cars, increased transit-oriented development, or increased global urbanization. There are managerial implications for such planners and government officials. For example, when ride-sharing platforms were initially emerging, many local planners and officials were wary of embracing them, unsure about the pros and cons. This research aims to provide more analysis into the potential pros and cons on the urban environment from existing and future transportation changes.

This is a challenging set of issues to study for a few reasons. First is availability of appropriate data, second is identifying an appropriate causal estimation of effects. To date, discussion around the impact of the entry of these ride-sharing services has not been able to study individual-level analysis or heterogeneity, instead studying the impact at the city level, such as aggregate entrepreneurship, drunk-driving, or traffic levels [2, 8, 12]. Furthermore, using data mining techniques to identify patterns of local consumption data may not provide useful causal insights for policy counterfactual analysis. Such patterns could be at risk of being influenced by city-specific unobservable local or national time trends, or self-selection bias of consumers.

Our motivating hypothesized mechanism is that mobile-app-enabled transportation services enable easier access to transportation — rather than just a replacement for existing transportation — which leads to lower transportation frictions for adopting consumers. Following from these lower frictions, we can expect several possible effects from the entry of Uber/Lyft. In this paper's analysis, we focus on the hypothesis that lower frictions lead to increased individual consumption with local restaurants and bars. If this hypothesis were not true, we would expect no changes in the consumption behavior of adopting consumers (relative to a comparable reference group), despite observing high usage of Uber/Lyft, since it would simply be replacing existing consumer transportation.

In this study, we use a novel and fine-grained panel dataset to estimate how such peer-to-peer transportation services might affect individual urban behaviors. We investigate the impact of the entry of the UberX and Lyft ride-sharing platforms. Our unique contributions will be taking advantage of a natural experiment of the entry of peer-to-peer sharing platforms across time and locations, and linking that with longitudinal, detailed data sources of consumer transactions and information. This estimation is further strengthened by controlling for potentially endogenous entry of Uber/Lyft

KDD '17, August 13-17, 2017, Halifax, NS, Canada

978-1-4503-4887-4/17/08…$15.00
DOI: 10.1145/3097983.3098058

(location-specific trends) and self-selection of adopters, by identifying an appropriately comparable non-adopter reference group. This enables a difference-in-differences analysis with propensity score matching to estimate the treatment effect on the treated, i.e. how adopters of Uber/Lyft may change their consumption patterns relative to a similar, but non-adopter reference group. This method of studying of the effect of adoption on adopters is similar to recent literature [17].

The fine-grained panel data we use are based on fine-grained individual credit card transactions from 2012-2016, from a total of over 40K continuously active accounts and 25M transactions. This range includes time periods both before and after the availability of Uber/Lyft. In our main findings, we identify, for the post-Uber/Lyft-entry period in various locations: first, an increase in the frequency and total amount spent on local food/drink businesses, but no notable increase in the amount spent per food/drink transaction. Second, we identify a strong relationship between an individual's usage intensity of Uber or Lyft and increases in food/drink frequency/spending. This is consistent with our hypothesis that the proposed mechanism is the usage of these P2P transportation platforms that leads to reduced local consumption frictions, more so than just the availability or adoption.

1.1 Related Literature

Peer-to-peer ride-sharing services, such as Uber and Lyft, are large companies associated with the broader "sharing economy" — AirBnB with lodging, Uber, Lyft, Didi or BlaBlaCar with transportation, TaskRabbit with freelance labor, etc [16]. According to a recent study by Pew Research Center in May 2016[1], 72% of U.S. population have engaged with such shared or on-demand services. A useful stream of recent research has provided empirical evidence on the impact of the entry of these "sharing economy" platforms. Based on the growth of various sharing economy platforms and companies, there have been several recent studies analyzing the impacts of the sharing economy. Several papers have focused on impacts on competitors within the related industry. In transportation, Uber data has been used to analyze New York City transportation patterns as well as corresponding relationship with taxi competition[2]. Uber's internal research team has also studied Uber's impact on increasing transportation availability in New York City [9]. [18] find a reduced revenue effect of AirBnB entry on the Texas hotel industry by using a natural experiment approach. By using a structural econometric approach, [5] conducted demand estimation and also found a reduced revenue impact of AirBnB growth on the national hotel industry.

Other papers have studied how the sharing economy has impacted complementary or indirectly-related sectors. [8] studied the entry of the Uber platform's effect in a societal context, using the varied entry times of Uber to estimate a reduction in alcohol-related motor vehicle homicides. [2] studied the impact of Uber entry on the volume of campaign launches on Kickstarter to examine the local entrepreneurial activity. [12] and [14] also studied the impact of Uber on aggregate local traffic and debt, respectively. A common thread for all of the above literature is looking at aggregate impacts of the sharing economy, rather than individual-level changes. Other theoretical literature on the impact of indirect impact of the sharing economy focus on how the impact on manufacturers of the shared goods may be affected by the entry peer-to-peer rental markets [1, 11]. Regarding the sharing economy itself, [4, 7, 10] focus theoretically on the management or preferences of the sharing economy companies, platforms, or participating consumers.

More broadly, outside of the sharing economy itself, this paper also contributes to a growing literature on the societal and offline impacts of changes in information/digital technology. In this vein, papers have looked at the entry of Craigslist had impacts beyond on the classified ads sector [15], but also on local behavior. [3] and [13] look at the impact of the availability of Craigslist on HIV transmission and prostitution trends, respectively. [6] study generally the growth of Internet investment and local wages. Methodologically, we reference a literature of using matching techniques to perform difference-in-differences analyses on adoption effects. In our dataset, we identify adopters, before and after an exogenous shock of Uber/Lyft availability. To ensure our comparison of adopters to non-adopters is more apt, we use matching techniques to identify a closer comparison group of non-adopters, similar to [17], in their study of iPad app adoption on purchases.

2 FINE-GRAINED LONGITUDINAL DATA

From a large U.S. financial institution, we have access to all customer-level transactions from a random, anonymized national sample of individuals across over 8 years (100 months). It is over-sampled in two particular metropolitan areas of interest to the institution. Several sources of spending are available; in this analysis we focus on the credit card consumers and transactions from January 2012 through May 2016, about four and a half years of observations. This date range is an appropriate data range to use because while both UberX and Lyft (popular peer-to-peer transportation platforms) were both introduced in San Francisco in mid-2012, their national expansion and growth mostly occurred in 2013 and 2014, as seen in Figure 1 for several metropolitan areas, and the large majority of our observed consumers are not from San Francisco.

For each transaction, we observe the account ID, day of transaction, amount spent, the standardized SIC store categorization, and importantly, the name of the store[3]. Debit, prepaid card, and ATM transactions also have more specific zip code, city, and address information of the store. With information on each specific store's merchant name, we can identify individual Uber and Lyft transactions.

While the data contains many observed transactions and consumers, our focus of interest is on consumers in metropolitan areas that received the Uber/Lyft entry. This enables study of the effect of entry and adoption by having enough consumers in each area to estimate city-specific, counterfactual time trends. Additionally, to avoid possible confounding in the data from exogenous changes in consumption, we want to only consider consumers who did not join

[1] http://www.pewinternet.org/2016/05/19/the-new-digital-economy/

[2] https://fivethirtyeight.com/features/uber-is-taking-millions-of-manhattan-rides-away-from-taxis/

[3] Any stores with fewer than 10 transactions will be truncated to ensure it does not have personally identifiable information. For example, in some credit card transactions, the transaction number will be described in the store name. These unique numbers are scrubbed from our dataset to ensure privacy, but the more general store information, if no longer identifiable, is still kept.

or leave their credit card account between 2012 and 2016. Lastly, we want to ensure a minimum level of activity and transaction frequency to ensure that we are only comparing consumers who are active users of their credit card. These constraints are important to ensure that we appropriately identify patterns of consumer spending both before and after the availability of Uber/Lyft.

To do so, we define a "core sample" of the consumers in the data. We identify each consumer's mailing address city in each month. We remove consumers who were in multiple cities during 2012 - 2016 to again reduce potential confounding. Having connected each consumer to a city, we then identify Uber/Lyft transactions using the transaction merchant names (described below) and identify the top 30 cities based on total number of Uber/Lyft transactions. These 30 cities correspond to 12 broader metropolitan areas. We keep only consumers living in these 30 cities because we want to study the effect of Uber/Lyft on adopters. If the adoption and usage of Uber/Lyft is small in a city, then this city is not as useful in estimating an effect. Lastly, we also ensure a minimum of at least 10 transactions in each complete year of our sample. This results in 14% of the original number of consumers in the data. With this reduced core sample, our analysis is representative for a sample of non-moving consumers, in cities with notable Uber/Lyft adoption, who already had their credit card accounts in 2012, and maintained spending activity in each year of 2012-2016.

While this reduction of our data into a "core sample" reduces the amount of data substantially, for this analysis, we prefer this smaller sample. The data in this sample is more trustworthy against risks from possible external confounding influences. The core sample still has enough data to use for the main analysis and in our robustness and heterogeneity analysis, we relax the definitions of our sample of interest.

2.0.1 Measuring Uber/Lyft Usage. We can identify whether each transaction is an Uber/Lyft transaction. We take all of our transactions and keep only those that specifically match certain Uber merchant name strings (identified by manually looking at all occurring strings with the string "UBER" in the merchant name and removing false positive merchant names). We also do the same for the "LYFT" string, which has no confounding occurrences. For the cities where we observe the most consumers, we plot the usage of Uber/Lyft per consumer below in Figure 1. The median number of trips amongst adopters in our core sample is 7 total trips, with a median price of $13.70. Approximately 8% of our observed unique accounts are observed to have paid for Uber/Lyft in our time sample.

2.0.2 Measuring Outcomes of Interest. Our first outcome of interest is an individual's number of transactions, in a week, at food/drink merchants. This follows from our proposed mechanism: lower transportation frictions will lead consumers to engage more frequently with local businesses. We further hypothesize that this effect from lower frictions will be stronger for leisure activities, such as going out to eat or drink. Thus, we keep only transactions for merchants that have Standard Industrial Classification (SIC) (which describes the store type) of 5812 (Retail-Eating Places) or 5813 (Retail-Drinking Places). These are considered food/drink transactions. There is an average of 0.7 food/drink transactions per week per account.

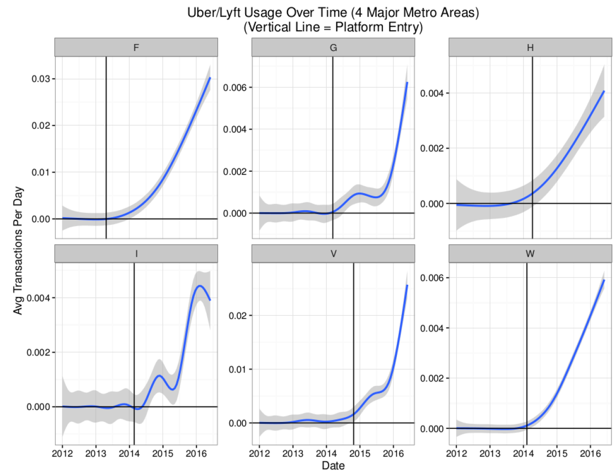

Figure 1: Showing the Uber/Lyft transactions per day, before and after the entry of Uber/Lyft, in 6 major metropolitan areas. On the bottom-left, in City I, the fluctuation in usage reflects a shutdown and re-opening of Uber/Lyft in that locale.

Correspondingly, a second outcome variable of interest is the total amount spent by an individual, in a week, at food/drink merchants. We expect that if the number of transactions increases, the total amount spent should also increase. However, when we consider a third outcome variable: the average amount spent by an individual, in a week, at food/drink merchants; we do not expect this to increase, because we hypothesize that frictions should not notably affect the price range of chosen merchants. There is an average of $33 spent on food/drink transactions per week per account.

Additionally, we are also interested in the geographical nature of the Uber/Lyft treatment. As we discussed, for consumers in a city, lower transportation frictions could ease access to some merchants or neighborhoods. Some merchants may be in difficult-to-access places, such as far away from a subway or public transit stop, having limited parking, or simply being further away from a consumer's home. Detailed geographic locations are unavailable, so as a simple measure and our fourth outcome variable of interest, we measure the number of unique food/drink merchants an individual visits, in a *month*. There is an average of 11 unique food/drink merchants per month per account.

Lastly, for each of the outcome variables, we use the natural-log of each variable plus 1 as the outcome of interest. This is because each of the four outcome variables are skewed. This also allows us to estimate the elasticity of our outcome variables with respect to the predictive features. Also, we identify robustness in the qualitative results by aggregating them at different time windows: {day, half-week, week, and month}.

3 ECONOMETRIC MODELING FOR CAUSAL ANALYSIS

In our main estimation, we use a difference-in-differences style approach to estimate the treatment effect entry of ride-sharing services on our outcomes of interest for the Uber/Lyft adopters, but keeping only a limited matched control group. As described above, these outcomes include individual spending frequency, total amount spent, amount spent per transaction, and number of unique merchants. To do so, we adopt a similar difference-in-differences approach as previous literature studying adoption effects (e.g., study of iPad app adoption by [17]).

Our key identification first comes from the variation of the times of the entry of Uber and Lyft services into different U.S. metropolitan areas. This variation enables identification between the potential impact of entry versus national time trends. We argue that this entry date is exogenous to individuals, suggesting a natural experiment on adopters — i.e., we observe adopters who were initially unable to adopt and use the service, until an exogenous outcome enabled a new service for such consumers. Combined with controlling for individual fixed effects (identified by pre-entry data), this potentially exogenous entry supports identification of a relationship between outcomes and a constant effect in the post-entry period.

However, there may be local-specific time-trends, or correlation between entry and local-specific unobservables. Additionally, consumers who do not adopt soon after entry may be susceptible to individual-specific time-varying unobservables, such as changes in income. To control for such local time trends and the self-selection of the non-immediate adopters, we use propensity score matching to identify an appropriate, similar reference group of non-adopters to compare against the adopters of interest. These consumers are matched on their propensity to adopt, using various features, including consumption behavior and locale. This enables a difference-in-differences panel data analysis, at the local city level. As we show in Figure 2, matching also improves the parallel trends assumption pre-entry at the individual level.

If our hypothesis is incorrect and the alternate explanation is that the Uber/Lyft services simply just replace existing transportation choices — e.g., driving or taxis — we should not expect to see positive associations between Uber/Lyft entry or usage and our outcomes of interest. Since we show that the adopters are similar to the non-adopters in their consumption behavior before adoption, if Uber/Lyft have no impact on consumption behavior, then consumption should also remain similar after the entry and usage of Uber/Lyft.

Importantly, in our main estimation, we use the sample of adopters and matched reference group. We keep all the consumers who we identified as adopters, having paid for Uber/Lyft (n = 664), and a corresponding propensity score matched sample of non-adopters (n = 664), with a similar estimated propensity to adopt. These two sets of consumers and their transactions are the matched group and data. The propensity score matching procedure is described in more detail below.

Without matching, the adopters appear to be quite different on average than the non-adopters, perhaps due to many non-adopters not utilizing their credit card frequently. However, once we only consider the propensity matched group, there appears to be a strong parallel trend between the adopters and non-adopters in multiple metropolitan areas. This parallel trend is consistent with our difference-in-differences hypothesis that the matched reference group provides an estimate of the counterfactual consumption outcomes for the adoption group, including time trend effects pre-entry. Of course, adoption is still a choice by consumers, which is why we emphasize that, unlike a policy analysis, we focus on potential effects on just those treated. (Potential spillover effects of Uber/Lyft on non-adopters are possible, but captured by the reference group in this analysis.)

With this matched data, we perform the following difference-in-differences style estimation. We control for time-invariant, individual-level differences using account-level fixed effects. We also include city-level fixed effects and time dummies to control for seasonality in our outcome variables.

Importantly, since financial activity is varying over time and may be correlated with both Uber/Lyft and our outcome variables, we want to avoid confounding our estimates with periods in time with higher spending. Thus, as a control, for each corresponding month, we include each individual's total number of transactions, spending, and average spent per transaction. Lastly, we also use clustered robust standard errors at the city-level in case there are city-level correlations in errors among consumers. This results in the following estimating equation:

$$Y_{ijt} = \alpha_i + \lambda_j + \tau_t + \beta X_{it} + \nu E_{jt} + \gamma I_i E_{jt} + \varepsilon_{ijt}$$

where Y_{ijt} are an individual i's observed outcomes in time period t and location t, X_{it} are observed individual time-varying characteristics (e.g., monthly total spending attributes), I_i is an indicator if an individual is an adopter of Uber/Lyft, and E_{jt} is an indicator if location j has Uber or Lyft services available. In this case, γ is our coefficient of interest, for the post-entry period treatment effect.

We focus primarily on the post-entry period because, due to network effects, an adopter may be exposed to the impact of Uber/Lyft even before their first trip. For example, they may share a ride with friends, or friends of our adopter sample who also utilize Uber/Lyft may encourage food/drink consumption on our adopter sample. Additionally, the post-entry period has a simpler identification, due to the exogenous availability of Uber/Lyft — though estimating long-term impacts is more difficult.

Next, we introduce the key aspect of interest, measuring individual actual usage of Uber/Lyft, our specific hypothesized mechanism of the effect. As a reminder, we propose that Uber/Lyft lower transportation frictions for some consumers, enabling them to engage more with local businesses. We introduce Q_{it}, which measures both the number of and amount spent on Uber/Lyft transactions for an individual i in time t. In this case, we expect if actual usage is observed, this should drive the associations we observe, whereas after controlling for usage, Uber/Lyft entry or adoption should not have a direct associated relationship with outcomes.

An important note is that usage may be correlated with an individual consumer being in a period of high consumption. The issue is that this would only be observed for adopters, while we cannot observe if non-adopters are in a high period. Thus, the estimate for the association of Uber/Lyft usage with outcomes could include both the influence from usage itself (mechanism) along with influence from being more likely in a high consumption time period. To address this concern, we control for individual-specific spending trends in all categories (X_{it}), and in Table 3, we vary the time window of analysis, expecting reduced confounding when the

Figure 2: Average per day number of food/drink transactions between adopters and non-adopters for two cities with a large bank market share. On the top is the entire core (unmatched) sample, in the middle and bottom are two cities using only the matched reference group.

window is at the month-level. (As seen in the analysis of Table 2, there does not appear to be confounding influence from potential high-consumption periods in the number of unique merchants per month).

$$Y_{ijt} = \alpha_i + \lambda_j + \tau_t + \beta X_{it} + \nu E_{jt} + \gamma I_i E_{jt} + \eta Q_{it} + \varepsilon_{ijt}$$

In a third specification, we also include an indicator, A_{it}, for the post-adoption period for each individual. We include this variable to identify if there are Uber/Lyft adoption effects in addition to the entry effects. In fact, we do not expect there to be adoption effects because our proposed mechanism is derived from actual Uber/Lyft usage, not merely a one-time adoption. Importantly, in order to appropriately estimate the association of adoption with outcomes, for the reference sample, we simulate their adoption date (i.e. date of first use) to their propensity score matched control consumer (as also utilized in [17]). By doing so, we ensure that each adopter has the same post-entry period as their corresponding control consumer. Thus, we use the following estimating equation:

$$Y_{ijt} = \alpha_i + \lambda_j + \tau_t + \beta X_{it} + \nu E_{jt} + \gamma I_i E_{jt} + \zeta A_{it} + \psi I_i A_{it} + \eta Q_{it} + \varepsilon_{ijt}$$

3.0.1 Matching Procedure. In this analysis, we use a nearest-neighbor static propensity score matching. A concern about analyzing the entry of Uber/Lyft is that it could be correlated with city-specific trends, such that time trends in other cities may not adequately control for. To better address this concern, we use propensity score matching to identify a comparable control sample. Additionally, as shown in Figure 2, the consumption behavior, before entry, of all non-adopters is not similar to those of the adopters. It's also possible that those who adopt on their credit card account are more likely to be individuals who use this bank's card as a "main" credit card, while the non-adopters are less likely to do so. To address these various concerns, we limit our control group using matching, including using consumption behavior as a predictor, on adoption propensity. This protects against the risk that the reference control group are simply those users are not "main" users of this bank. Further, if the non-adopter control group also is influenced by network effects of peers using Uber/Lyft, or using such services on other financial accounts, this would only serve to make our estimates more conservative.

We estimate a logistic regression on whether a consumer was identified as having adopted Uber/Lyft or not. As predictors for adoption propensity, we use the following individual-level features (only including data up until the date of Uber/Lyft entry though): {*city of consumer, zip code of consumer, number of months with transactions, mean number of total transactions per month, mean amount spent per month, mean amount spent per transaction per month, mean credit score from 2012 - 2016, and gender*}. In Figure 3, we show the distribution of adoption propensity for the resulting matched group (664 adopters, 664 matched non-adopters). This simple nearest neighbor matching seems to have identified a closely matched control group, with good support for the adopter group. Subsequently, Figure 2 also shows approximate parallel trends in food/drink consumption amongst the matched group. In this main analysis, we only used a simple matching process; in our extensions, we would like to improve the robustness of our matching using both different matching techniques and additional predictive features.

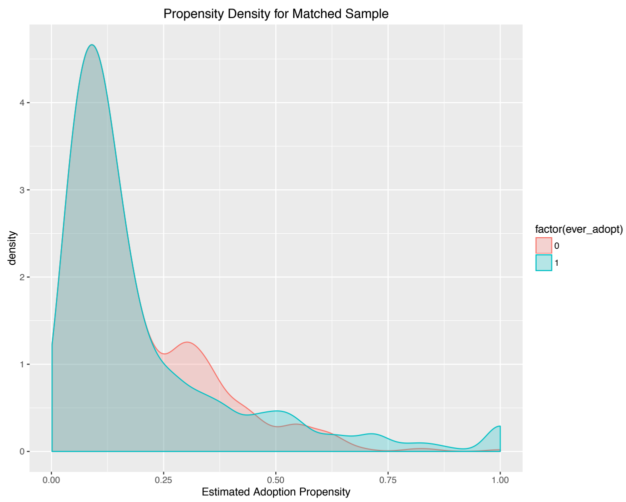

Figure 3: In the nearest-neighbor matched sample, with equal number of adopters and non-adopters, a plot comparing the similar, overlapping propensity density of the two groups. This shows the good support of the control group over the range of the adopter propensity.

4 MAIN FINDINGS

In our first result of interest, we focus on the number of transactions as the outcome variable and individual transactions aggregated to the week-level. These results are in Table 1 and illustrative of the results in the other tables. In column (1), the simple analysis shows a significant positive association for adopters, relative to the trends of the reference group, between the post-entry period and food/drink transactions. This is controlling for national week time dummies, city-specific (not metro area) and individual fixed effects, and using city-clustered standard errors. It suggests an associated 2% more total transactions per week for adopters in the entire multiple years of the post-entry period.

Next, we directly investigate our hypothesized mechanism that it is the use of Uber/Lyft that lowers local transportation frictions, enabling increased consumption. In column (2) of Table 1, when we control for an individual’s usage of Uber/Lyft usage in a week, we observe, relative to the reference group, a positive association between Uber/Lyft usage in that week and food/drink transactions for consumers. The post-entry period no longer has significantly positive association in this case. This is consistent with our hypothesized mechanism, since we would not expect an increase in transactions when we observe zero Uber/Lyft transactions — only when an individual utilizes the friction-reducing mechanism. Column (2) suggests an associated 7% increase in transactions per week for individuals per observed Uber/Lyft transaction.

As discussed in the methodology section, we recognize however that, if consumers have variability in spending frequency across time periods, this estimate could be confounded by the (unobserved) variable of being in a high frequency period. To control for this to some extent, we include controls for each user’s total monthly spending frequency, amount, and amount spent per transaction on all transaction types. Additionally, as discussed below, column (4) of Table 2 and column (5) of Table 3 also reduce this confounding concern: for example, when considering monthly transactions instead of weekly, we continue to see positive associations between monthly Uber/Lyft usage and food/drink transactions relative to the reference group.

In column (3) of Table 1, we also include controls for each adopter’s and their matched non-adopter’s post-adoption period. The comparison between the post-entry period and the post-adoption period is that the post-adoption period occurs, by definition, after the entry of Uber/Lyft, and can substantially later for some consumers (e.g., 1-2 years after entry for some). The results of column (3) actually show a negative association between the post-adoption period for adopters and consumption frequency, statistically netting out the initial positive association of the post-entry period for adopters. This is notable, consistent with the hypothesis that there could be an initial increase in consumption frequency post-entry for adopters, even without direct Uber/Lyft usage, perhaps driven by network effects. In later periods though, this indirect increase in consumption associated with even no Uber/Lyft usage may go away — instead, positive increases in consumption being associated only with the mechanism of interest, the usage of Uber/Lyft.

In Table 2, we study how the entry of Uber/Lyft is related to different outcome variables. The regression specifications and results of column (1) are the same as those in column (3) in Table 1, but with different outcome variables. In the second column, there is a notable positive association for adopters in the post-entry period and Uber/Lyft transactions for the total amount spent in each time period. This is consistent with our proposed hypothesis. However, for the amount-per-transaction outcome and the number of unique merchants outcome — aggregated at the month level — we see positive associations, but not significantly so. In our proposed mechanism, we do not expect the amount spent per transaction to increase, so this is consistent.

It is of note that while there is a significant positive association between Uber/Lyft transactions and transactions/amount spent, there is no such positive association with the number of unique merchants nor amount spent. This reduces the concern that estimates of Uber/Lyft weekly transactions and consumption are confounded by the presence of high frequency consumption periods. Otherwise, in high consumption periods, consumers could be considered to spend more and in more places, which we do not observe.

Further, we provide additional robustness checks in Table 3 by using several time windows to aggregate individual transactions. Each time window shows similar qualitative patterns. Note that, in Table 3, at the month level, we do not have a significantly positive associations of Uber/Lyft activity with the number of transactions per month. However, when we limit our scope to metropolitan area “A”, with the most accounts in the data (*70%* of the accounts), we do identify a significant positive associated 2.3% increase in transactions at the month-level. This suggests that further work can be done to assess heterogeneous effects between various cities. It also reduces the risk that the positive associations between Uber/Lyft transactions and food/drink transactions is driven only by high-frequency periods; while this association is strongest of course at the day-level, we expect that at the month level, variability in spending between months will be reduced, especially when already controlling for total monthly spending.

(food/drink transactions only)	Outcome: ln(number of transactions+1)		
	(1)	(2)	(3)
post_entry	−0.014	0.008	−0.018
	(0.011)	(0.011)	(0.012)
post_entry_for_adopter	0.017*	0.009	0.026***
	(0.008)	(0.007)	(0.007)
uber/lyft_transactions		0.069***	0.070***
		(0.008)	(0.008)
uber/lyft_spending		0.001***	0.001***
		(0.0003)	(0.0003)
post_adoption			0.013**
			(0.005)
post_adoption_for_adopter			−0.039**
			(0.007)
Monthly spending behavior controls	Yes	Yes	Yes
Observations	292,891	292,891	292,891
Time Dummies	specific week	specific week	specific week
Fixed Effects	city, account	city, account	city, account
Clustered Errors	city-level	city-level	city-level

Table 1: Regression of Uber/Lyft entry on the number of food/drink transactions, aggregated at the individual-week level. Notable is the relationship between Uber/Lyft usage in the week and transactions in the week.

(food/drink transactions only)	ln(# of transactions)	ln(amount spent)	ln(amount per transaction)	ln(# unique merchants)
	(1)	(2)	(3)	(4)
post_entry	−0.018	−0.040	−0.029	−0.035
	(0.012)	(0.039)	(0.022)	(0.022)
post_entry_for_adopter	0.026***	0.095***	0.011	0.019
	(0.007)	(0.027)	(0.018)	(0.014)
uber/lyft_transactions	0.070***	0.234***	0.005	−0.004
	(0.008)	(0.035)	(0.006)	(0.003)
uber/lyft_spending	0.001***	0.004***	0.0003	0.00004
	(0.0003)	(0.002)	(0.0004)	(0.0003)
post_adoption	0.013**	0.040**	−0.009	−0.013
	(0.005)	(0.019)	(0.020)	(0.014)
post_adoption_for_adopter	−0.039**	−0.142***	−0.006	−0.006
	(0.007)	(0.023)	(0.018)	(0.015)
Monthly spending behavior controls	Yes	Yes	Yes	Yes
Observations	292,891	292,891	292,891	67,787
Time Dummies	specific week	specific week	specific week	specific month
Fixed Effects	city, account	city, account	city, account	city, account
Clustered Errors	city-level	city-level	city-level	city-level

Table 2: Comparison of week-level regression results across 4 different outcome variables (using regression specification 3 in Table 1). Notable is that transaction frequency and amount spent per time period have a positive relationship with Uber/Lyft activity, but not amount spent per transaction.

Additionally, while not presented here, we have also utilized Poisson and negative-binomial regressions as a robustness check and identified similar analysis results for the post-entry period, post-adoption period, and Uber/Lyft usage.

Lastly, in Table 4, we present some initial analysis on temporal heterogeneity in the impact of Uber/Lyft entry. We identify a larger positive association between Uber/Lyft transactions and food/drink transactions on Fridays, Saturdays, Sundays, and Mondays (perhaps

(food/drink transactions only)	by day	by half-week	by week	by month	by month (metro area A)
	(1)	(2)	(3)	(4)	(4)
post_entry	-0.015^{*}	-0.005	-0.018	-0.024	
	(0.009)	(0.006)	(0.012)	(0.026)	
post_entry_for_adopter	0.014^{**}	0.014^{***}	0.026^{***}	0.064^{***}	0.046^{***}
	(0.005)	(0.004)	(0.007)	(0.018)	(0.012)
uber/lyft_transactions	0.594^{***}	0.108^{***}	0.070^{***}	0.010	0.023^{***}
	(0.052)	(0.010)	(0.008)	(0.007)	(0.004)
uber/lyft_spending	0.003	0.001^{***}	0.001^{***}	0.001	-0.0002
	(0.002)	(0.003)	(0.0003)	(0.001)	(0.0003)
post_adoption	0.008^{**}	0.006^{**}	0.013^{**}	0.025	0.005
	(0.003)	(0.003)	(0.005)	(0.017)	(0.013)
post_adoption_for_adopter	-0.019^{***}	-0.022^{***}	-0.039^{**}	-0.076^{***}	-0.065^{***}
	(0.006)	(0.023)	(0.007)	(0.016)	(0.016)
Monthly spending behavior controls	Yes	Yes	Yes	Yes	Yes
Observations	1,996,519	578,108	292,891	67,787	39,962
Time Dummies	specific day	specific half-week	specific week	specific month	specific month
Fixed Effects	city, account	city, account	city, account	city, account	city, account
Clustered Errors	city-level	city-level	city-level	city-level	none

Table 3: Comparison of heterogeneity in the relationship of Uber/Lyft usage and consumption totals, using various time aggregation windows.

when some Sunday transactions enter the data). Intuitively, this provides additional evidence consistent with our proposed mechanism that the availability of Uber/Lyft and lowered transportation frictions will have a stronger influence on leisure activities. Similarly, when we group the data by half-week (where a week is split into M-Th and F-Su), we the similar stronger association on weekends. We recognize that the magnitude of these numbers are over-estimated though, because daily aggregation is much more influenced by temporal variability in consumption — instead, the takeaway should be to note that there seem to be temporal differences in association of Uber/Lyft transactions and food/drink transactions.

5 CONCLUSION & FUTURE WORK

In this paper, we utilize a novel dataset to present a quasi-experimental examination of how the emerging growth of peer-to-peer car sharing services may have affected local consumer mobility and consumption patterns. This is a relevant policy and management question for cities. From the analysis, we identify, for the post-Uber/Lyft-entry period, first, an associated increase in the frequency and total amount spent on local *food/drink* businesses from Uber/Lyft entry and particularly from Uber/Lyft usage intensity, but no notable increase in the amount spent *per* food/drink transaction. More directly, consistent with our hypothesized mechanism of reduced local consumption frictions, if we consider the *usage* of these P2P transportation platforms, we see that this is what drives the observed association, moreso than the adoption or entry. This research provides additional empirical evidence of the impact of recent and potential future change in urban transportation for city officials.

For future work, we plan to investigate the heterogeneity in the impact of Uber/Lyft more closely amongst demographics, weather patterns, spatially, and temporally. We plan to identify fine-grained geo-locations of merchants, enabling a closer investigation into how Uber/Lyft may have shifted spatial movements, and incorporate other sources of financial data. To demonstrate the robustness of our proposed mechanism, we plan to analyze other categories of spending, either of interest, such as other modes of transportation, or online-related categories for falsification tests. We plan to zoom in a few metropolitan areas of interest to more closely assess how consumption patterns may have shifted. While our presented results are robust to various time windows and alternative assumptions for the regression function (e.g. Poisson), we plan to demonstrate robustness across various subsamples of the data, regression specifications (e.g. relative time periods, as used by [8]), and propensity matching techniques.

REFERENCES

[1] Vibhanshu Abhishek, Jose A Guajardo, and Zhe Zhang. 2016. Business Models in the Sharing Economy: Manufacturing Durable Goods in the Presence of Peer-To-Peer Rental Markets. *SSRN* (Jan. 2016).
[2] Gordon Burtch, Seth Carnahan, and Brad N Greenwood. 2016. Can You Gig it? An Empirical Examination of the Gig-Economy and Entrepreneurial Activity. *SSRN* (March 2016).
[3] Jason Chan and Anindya Ghose. 2014. Internet's Dirty Secret: Assessing the Impact of Online Intermediaries on HIV Transmission. *MIS Quarterly* 38, 4 (Dec. 2014), 955–A3.
[4] Liran Einav, Chiara Farronato, and Jonathan Levin. 2016. Peer-to-Peer Markets. *Annual Review of Economics* 8, 1 (Oct. 2016), 615–635.
[5] Chiara Farronato and Andrey Fradkin. 2016. Market Structure with the Entry of Peer-to-Peer Platforms: The Case of Hotels and Airbnb. *Working Paper* (2016).
[6] Chris Forman, Avi Goldfarb, and Shane Greenstein. 2012. The Internet and local wages: A puzzle. *The American Economic Review* 102, 1 (2012), 556–575.
[7] Samuel P Fraiberger and Arun Sundararajan. 2015. Peer-to-Peer Rental Markets in the Sharing Economy. *NYU Stern School of Business Research Paper* (2015).
[8] Brad N Greenwood and Sunil Wattal. 2017. Show Me the Way to Go Home: An Empirical Investigation of Ride-Sharing and Alcohol Related Motor Vehicle

(food/drink transactions only)	by day	by half-week
	(1)	(2)
Post-entry, post-adoption controls for adopters and non-adopters; and monthly behavior controls	Yes	Yes
uber/lyft_transactions:Monday	0.632^{***}	
	(0.057)	
uber/lyft_transactions:Tuesday	0.532^{***}	
	(0.076)	
uber/lyft_transactions:Wednesday	0.515^{***}	
	(0.036)	
uber/lyft_transactions:Thursday	0.563^{***}	
	(0.106)	
uber/lyft_transactions:Friday	0.649^{***}	
	(0.057)	
uber/lyft_transactions:Saturday	0.616^{***}	
	(0.079)	
uber/lyft_transactions:Sunday	0.598^{***}	
	(0.045)	
uber/lyft_transactions:M-Th		0.093^{***}
		(0.011)
uber/lyft_transactions:Fr-Su		0.117^{***}
		(0.012)
Observations	1,996,519	578,108
Controls	time, fixed effects, cluster robust errors	time, fixed effects, cluster robust errors

Table 4: Comparison of heterogeneity in the treatment effect of Uber/Lyft usage

Fatalities. *MIS Quarterly* 41, 1 (2017), 163–187.
[9] Jonathan Hall, Cory Kendrick, and Chris Nosko. 2015. The effects of Uber's surge pricing: A case study. (2015).
[10] John J Horton and Richard J Zeckhauser. 2016. *Owning, Using and Renting: Some Simple Economics of the "Sharing Economy".* Technical Report. National Bureau of Economic Research.
[11] Baojun Jiang and Lin Tian. 2016. Collaborative Consumption: Strategic and Economic Implications of Product Sharing. *Management Science* (Nov. 2016), mnsc.2016.2647.
[12] Ziru Li, Yili Hong, and Zhongju Zhang. 2016. Do Ride-Sharing Services Affect Traffic Congestion? An Empirical Study of Uber Entry. *SSRN* (2016).
[13] Probal Mojumder, Jason Chan, and Anindya Ghose. 2016. The Digital Sin City: An Empirical Study of Craigslist's Impact on Prostitution Trends. In *Proceedings of International Conference on Information Systems 2016.*
[14] Tingting Nian, Yuyuan Zhu, and Vijay Gurbaxani. 2016. Uber on: Until you are free of debt . In *Workshop on Information Systems and Economics 2016.*
[15] Robert Seamans and Feng Zhu. 2013. Responses to entry in multi-sided markets: The impact of Craigslist on local newspapers. *Management Science* 60, 2 (2013), 476–493.
[16] Arun Sundararajan. 2016. *The Sharing Economy.* MIT Press.
[17] Kaiquan Xu, Jason Chan, Anindya Ghose, and Sang Pil Han. 2016. Battle of the Channels: The Impact of Tablets on Digital Commerce. *Management Science* 63, 5 (May 2016), 1469–1492.
[18] Georgios Zervas, Davide Proserpio, and John W. Byers. 2017. The Rise of the Sharing Economy: Estimating the Impact of Airbnb on the Hotel Industry. *Journal of Marketing Research* (Jan. 2017). DOI: http://dx.doi.org/10.1509/jmr.15.0204