What Drives Pricing Behavior in Peer-to-Peer Markets? Evidence from the Carsharing Platform BlaBlaCar*

Mehdi Farajallah (Marsouin)

Robert G. Hammond (North Carolina State University)

Thierry Pénard (CREM, University of Rennes 1)

August 2016

Abstract: How are prices and market outcomes determined on peer-to-peer platforms? More importantly, how should we expect price-setting and demand behavior to change as these markets mature? We provide the first empirical analysis of the world's leading carsharing platform, BlaBlaCar. Our econometric model explicitly accounts for the joint determination of price and quantity demanded and finds that pricing decisions evolve as drivers gain experience with the platform. More-experienced drivers set lower prices and, controlling for price, sell more seats. Our interpretation is that more-experienced drivers on BlaBlaCar learn to lower their prices as they gain experience. Further, we find that driver demographics matter. The demographic characteristic with the quantitatively largest effect is for drivers with an Arabic-sounding name, for whom there is meaningfully lower demand, despite the fact that these drivers set lower prices. In total, our results suggest that peer-to-peer markets such as BlaBlaCar share some characteristics with other types of peer-to-peer markets such as eBay but remain a unique and rich setting in which there are many new insights to be gained.

* We thank seminar and conference participants at the French Economic Association Congress, North Carolina State University, the Searle Center Conference on Internet Commerce and Innovation, and the 3rd Digital Economy Summer School, as well as Wendy Bradley and Franco Mariuzzo.

Introduction

The rise of the "sharing economy" and the success of sharing platforms like AirBnB and Uber have attracted the attention of economists and other academics as well as the popular press (Horton and Zeckhauser, 2016). Einav et al. (2016) define these platforms as peer-to-peer markets and emphasize their role in matching buyers to sellers to facilitate transactions with reduced scope for opportunistic behavior.[1] The traditional analysis of these types of markets focuses on markets such as eBay or AmazonMarketPlace; recently however, these analyses have begun to focus on markets for carsharing, lending, accommodation, home services, deliveries, or task assignments (Sundararajan, 2016).

We study the leading carsharing platform, BlaBlaCar, which is valued at $1.5 billion as of 2015.[2] BlaBlaCar connects a driver with empty seats to riders to share an intercity trip. The importance of BlaBlaCar has been emphasized by Sundararajan (2016), calling it "the company that dominates [the intercity carsharing] market" and noting that BlaBlaCar moves "as of 2015, more people every day than the US national rail system Amtrak" (Sundararajan, 2016 p. 12). We use this empirical setting to study the determinants of price setting and demand behavior in order to understand how we should expect these types of peer-to-peer markets to evolve moving forward. The uniqueness of the BlaBlaCar platform allows this study to offer particularly relevant insights for the policy discussions surrounding the sharing economy. To this end, we analyze of the joint determination of price and the quantity of seats demanded.

---

[1] Similarly, Rochet and Tirole (2006) and Evans et al. (2011) define these platforms as two-sided markets that bring together two groups of economic agents: sellers and buyers, hosts and guests, or drivers and riders.

[2] http://www.wsj.com/articles/blablacar-joins-ranks-of-billion-dollar-venture-backed-startups-1442433577

The BlaBlaCar setting is uniquely well suited for our study because its price-setting environment is different from other carsharing peer-to-peer markets. On platforms like Uber and Lyft, pricing is centralized by the market maker and thus price is the same for any driver offering a given trip at a given moment. In contrast, on BlaBlaCar, pricing is decentralized and drivers set their own price for each trip. This provides rich price variation and increases the scope for factors such as experience and social preferences to affect prices.

Our econometric model addresses the endogeneity of the driver's price and, controlling for price, models the number of seats sold. Our two sets of main results concern a driver's level of experience and demographic characteristics. First, we focus on how drivers' price-setting behavior evolves as they gain experience on the platform. The results suggest that more-experienced drivers set lower prices than less-experienced drivers; controlling for price, more-experienced drivers sell more seats. The price result is counter to evidence from other offline and online markets, where brand loyalty effects allow firms with more experience in the market to charge higher prices.

Our finding that drivers lower their prices as they gain experience suggests that nonpecuniary factors and learning are important in price setting on BlaBlaCar. Specifically, we find that drivers learn to set lower prices as they gain experience. Moreover, drivers' reputation (measured by the quantity and quality of ratings) has a quantitatively small effect on price. We conclude that prices and market outcomes on "sharing platforms" such as BlaBlaCar are determined differently than on other types of peer-to-peer markets such as eBay, in which reputation has been found to have a quantitatively large effect on pricing (Cabral, 2012).

Second, we find the demographic characteristics of a driver have strong predictive power for her price and the demand for her seats. The results suggest that female drivers sell more seats than male drivers. Matching drivers' first/given names to a database of names and their predominant country or region of origin, we classify drivers as having a French-sounding name, Arabic-sounding name, or neither. Based on this classification, drivers with a predominantly French name sell more seats, and drivers with a predominantly Arabic name sell fewer seats, relative to a control group of drivers with names that are neither predominantly French nor Arabic. An Arabic name is associated with a particularly strong negative effect on demand. Our proxy for country/region of origin is imperfect but, despite this measurement error, we find a disadvantage faced by drivers whose name sounds Arabic or Muslim.

Beyond the richness of the variation in price-setting behavior mentioned above, BlaBlaCar has additional features that make it particularly interesting to study. Most empirical studies of peer-to-peer market focus on settings where interactions are limited to online communication (e.g., eBay). In contrast, on BlaBlaCar, the interaction between the two parties begins online, but ends offline in the driver's car where they share a small space for up to a few hours together. Beyond the obvious pecuniary gains (revenue as a driver or savings as a passenger relative to alternative transportation), BlaBlaCar users may also be motivated by nonpecuniary factors and prosocial behavior (Benabou and Tirole, 2006; Gneezy, Meier, and Rey-Biel, 2011). For example, social and ecological motivations (i.e., to create social ties or reduce road congestion/gas emissions) may be important in determining supply and demand in these markets (Schor and Fitzmaurice, 2015).

Peer-to-peer markets have been argued to provide important efficiency gains (Edelman and Geradin, 2015; Horton and Zeckhauser, 2016). Specifically, these markets lower search and transaction costs (e.g., reduce information asymmetries) and allow fuller use of resources (e.g., increase car occupancy). Despite these benefits, these platforms have introduced regulatory issues, especially in the transportation and accommodation sectors. Many municipalities and countries have taken steps to restrict the use of residential property for short-term rental through AirBnB or ban nonprofessional drivers on Uber. Our contribution is to analyze pricing and demand behavior for a particularly interesting platform as users gain experience on it, which is essential to understanding how peer-to-peer markets should be expected to evolve over time. This evolution is a central part of the understanding that is required to formulate appropriate policy moving forward.

The literature on peer-to-peer markets is growing but many important questions remain open (Einav et al., 2016). In a study of AirBnB, Edelman and Luca (2014) show that rental prices depend on demographics of the host: prices for non-black hosts are approximately 12% higher than black hosts for an equivalent rental. The authors say this finding is suggestive of digital discrimination, which is related to our findings regarding Arabic names. Other work on peer-to-peer markets has also focused on AirBnB (e.g., Zervas et al., 2014). There is also a literature on the effect of online ratings and reputation in peer-to-peer markets (Bolton et al, 2013; Dellarocas, 2013; Cabral and Hortacsu, 2010; Houser and Wooders, 2006; Jin and Kato, 2006; Jolivet et al., 2016; Melnik and Alm, 2002; Resnick and Zeckhauser, 2002; Resnick et al. 2006). However, rating a product is quite different from rating a personal experience with a driver or a host (Zervas et al., 2015 and Fradkin et al., 2014).

The paper is organized as follows. The next section introduces the carsharing platform BlaBlaCar. Section 2 describes our data and Section 3 explains the empirical methodology to analyze market outcomes and address the simultaneity of price setting and demand. Section 4 presents the main econometric results and Section 5 concludes.

1. The carsharing platform BlaBlaCar

Founded in France in 2006, BlaBlaCar has become the leading carsharing platform.[3] BlaBlaCar offers intercity ridesharing services, connecting drivers with empty seats to people who are traveling on the same trip (see Figure 1 in Appendix A). Drivers earn money and passengers save on travel expenses (given that the typical trip on BlaBlaCar is cheaper than the corresponding train or bus ticket). As of 2016, BlaBlaCar operates in 22 countries (mainly in Europe, but also in Mexico, India, Russia, and Brazil). In April 2015, BlaBlaCar acquired the second largest European carsharing company carpooling.com, which expanded BlaBlaCar to more than 20 million members. The company has a valuation of $1.5 billion. BlaBlaCar has not seen the types of regulatory battles faced by carsharing companies like Uber, because BlaBlaCar is considered a not-for-profit ride service. The stated purpose of the money received by drivers is only to share the cost of the trip.

Over 2 million people use BlaBlaCar every month, around 29% of whom are drivers. The average BlaBlaCar user is 34 years old, with 14% of drivers and 36% of passengers being students. Registration on BlaBlaCar is free but passengers pay fees that are about

[3] The name BlaBlaCar comes from the French word "blabla" that is the English equivalent of blah. Driver profiles can display their "talking" preference: Bla if they do not like to talk with passengers, BlaBla if they like to talk a little, and BlaBlaBla if they like to talk a lot.

15% of the price of the ride paid to the drivers.[4] Similar to most other peer-to-peer markets, passengers and drivers are asked to rate each other and write reviews.

For each trip, BlaBlaCar suggests a "recommended price" based on the trip distance and the estimated price of fuel and tolls.[5] The recommended price does not depend on the number of seats offered. The driver is allowed to adjust the price up or down, with the minimum (maximum) price set as 50% (150%) of the recommended price.[6] In February 2012, BlaBlaCar introduced a price color classification, where the driver's price displays to potential riders as green if the driver chooses a price that does not exceed the recommended price. Otherwise, the price is orange (up to 125% of the recommended price) or red (between 125% and 150%).

2. Data collection

Data were collected from February 2013 to March 2014, but our analysis focuses on the period from August 2013 to March 2014, when data were collected daily. We selected 43 French intercity trips, chosen to ensure a representative sample of trips that are offered on BlaBlaCar and to generate variability in trip distance (e.g., short and long trips). Of these 43 trips, 40 trips were offered at least several times and our final sample contains these 40 trips. The shortest trip is Nimes-Montpellier (56 km) and the longest trip is Paris-Marseille (774 km). Trips from across France were selected: trip between provincial cities as well as trip between Paris and a provincial city. The list of trips is

[4] Booking fees and VAT are added to the price that the passenger pays to the driver. The fees earned by BlaBlaCar are composed of a fixed component (€ 0.89) and a variable component (9.90%, of the price paid to the driver). A VAT of 20% is added to these fees.
[5] For instance, in 2015, the recommended price of BlaBlaCar was automatically calculated as follows: .065 € per kilometer and per seat if the driver takes a toll road and .048 € per kilometer and per seat otherwise.
[6] https://www.blablacar.in/faq/question/how-do-i-set-my-price

available in the appendix, including descriptive statistics (distance, number of observations, and unique drivers for each trip). There are 41 unique cities (33 of more than 100,000 inhabitants and 8 of less than 100,000 inhabitants).

The data collection procedure was automated. For each trip, we collected all offers, resulting in 948,789 observations from 297,582 individual drivers. The data collection script scraped the BlaBlaCar website, resulting in multiple snapshots of each observation (e.g., three days before departure, two days, etc.). Here, we focus on the last observation for each offer. The data contain the departure and arrival cities; departure date and hour; driver name; and their profile (gender, age, etc.); whether the driver's photo is shown; and declared preferences for smoking, pets, music, and talking (dislikes talking, likes a little talking, or likes a lot of talking). For each trip, we have the number of seats available, the price, and price color (green, orange, or red).

When these data were collected, the rating mechanism of BlaBlaCar allowed only a positive or negative rating. In our data set, members' reputation is, therefore, measured by the number and percentage of positive ratings that were received.[7] Controlling for reputation, a driver's five-level status measures her experience level. This status is publicly observable on the drivers' profile. A driver is classified as newcomer, intermediate, experienced, expert, or ambassador, as shown below.

	Newcomer	**Intermediate**	**Experienced**	**Expert**	**Ambassador**
Profile completion		> 60%	> 70%	> 80%	> 90%
Number of ratings		1 rating	3 ratings	6 ratings	12 ratings
% positive ratings		> 60%	> 70%	> 80%	> 90%
Seniority		1 month	3 months	6 months	12 months

[7] As of 2015, BlaBlaCar's rating system has expanded to include a five point scale.

Finally, we also have the comfort of the driver's car, where higher values indicate a more luxurious vehicle. Comfort is self-classified by the driver and 11.36% of drivers do not disclose any comfort level.

Table 1 presents summary statistics. Panel A summarizes the outcomes of interest for the entire sample, while Panel B presents our explanatory variables, including each driver only once. Regarding drivers' experience, the number of ratings received is 8.4 on average, while driver status is 2.7 and car class 2.3 on average.

Drivers are 36 years old, on average, and around 40% of listings are by females. Using a database of names and associated country/region of origin, we classify drivers by the most prominent region of origin of their first/given name.[8] Our approach classifies 67% of drivers as having a French-sounding name (e.g., Guillaume, Pierre, and Sophie), 5% as having an Arabic- or Muslim-souding name (e.g., Ahmed, Mariama, and Youssef), and the remaining 28% as having a name that is neither predominantly French nor Arabic (e.g., Kim, Mickael, and Tony).

Further, drivers include a picture with their profile in 39% of cases. 56% of drivers indicate that they play music during the trip, 9% allow pets, and 7% allow smoking. Finally, 26% of drivers offer roundtrip travel, while 12% of drivers allow a seat to be sold only after manually confirming the sale (as opposed to an immediate sale). Using these data, we estimate the econometric model that is explained next.

[Insert Table 1]

[8] We collected data on names origin from three web sources: www.insee.fr (the French National Institute of Statistics and Economic Studies) and two well-known French websites: http://www.prenoms.com and http://www.signification-prenom.net. The resulting database contained over 69,000 first/given names along with their country of origin, which was cross-checked across these three websites for accuracy.

3. Econometric Model

To understand the functioning of the BlaBlaCar carsharing platform, we perform a regression analysis to explain the prices charged and quantities sold by drivers in the data described above. We use data listed on BlaBlaCar starting from August 2013 until March 2014. We cleaned the data to exclude listings with a departure date after December 2014. The resulting data set contains 948,789 listings from 297,582 drivers.

Our econometric model uses a fixed-effects panel-data regression, with trip fixed effects. A trip is defined as a departure city-arrival city pair. There are 40 trips in the data, with the following five trips as the most commonly offered trips in descending order: (Nimes, Montpellier), (Nantes, Rennes), (Lille, Paris), (Lyon, Paris), and (Toulouse, Bordeaux). Including trip fixed effects allows us to control for the general characteristics of the trip, then look separately at specific factors that affect drivers' prices and riders' demand.

$$q_{ijt} = p_{ijt} + X_i + Z_j + W_t + \mu_j + e_{ijt}$$

with i=driver, j=trip and t=departure date.

Measuring the quantity sold

To define quantity sold (q_{ijt}), we use the panel nature of our data with repeated listings for a given driver. The data-extraction software used to gather data regularly visited hundreds of thousands of BlaBlaCar listing pages but instantaneous data collection is infeasible. As a result, the data occasionally contain a number of seats available variable that already reflects a lower quantity supplied than the true quantity supplied. That is, when the software scraped a given listing's page, the number of seats available may already be lower by one seat if a rider purchased a seat before the page was first scraped. While the data collection may miss a seat sold or two for a driver on a given listing, it is unlikely to systematically miss seats sold on all listings that a driver ever offers. As such, we construct a variable that is equal to the maximum number of seats available ever observed by the driver across all of her listings (seats in car).[9]

We use the number of seats in each driver's car to construct two quantity sold variables. First, fraction sold measures the proportion of seats that sold for the listing, varying between zero and one. Second, all seats sold is a dummy variable that equals one when fraction sold equals one, that is, when the number of seats available equals zero at the close of the listing. Importantly, the all seats sold dummy is robust to our approach for measuring the number of seats in the car. In particular, while we imperfectly observe the number of seats initially offered, we perfectly observe the number of seats available for each listing, irrespective of how soon or how often the data-extraction software gathered data on a listing. If zero seats are available when the listing closes, then all seats sold, by definition. The two quantity sold measures provide similar results in what follows, providing support for our approach for defining quantity sold.

Controlling for Price Endogeneity

The outcomes of interest are price and quantity sold. We present an instrumental variables regression analysis, where price is considered an endogenous variable that affects quantity sold. As such, we use an instrument that we argue affects the driver's price but has no effect on riders, except through its influence on price. The instrument is constructed from the trip-level panel nature of our data. Specifically, we link drivers

[9] While this measure could be noisy if a driver uses different cars for different trips, we have characteristics that help to identify the car used in a listing (specifically, the comfort level of the car). We construct an alternative measure of the maximum number of seats available, finding that the two measures largely overlap.

who offer a given trip to other trips offered by the same driver to construct the universe of listings offered by the driver during our sample period.

Importantly, this approach requires that we precisely identify drivers. Unfortunately, unlike other online marketplaces, BlaBlaCar does not use unique user IDs as part of its listing interface, an approach that is useful with eBay data, for example. As a result, we need to identify drivers as carefully as possible to identify which listings were offered by the same driver. To do so, we use three variables in our data: name, age, and gender. It is important to note that we have name information on the driver's first/given name as well as the first initial of the driver's last name.[10] Coupled with age and gender, we are able to classify drivers with a high degree of precision.

Having identified drivers, we construct driver characteristics in three ways: over all trips, over the trip in question, and over all trips other than the trip in question. To be clear, this implies that if a driver is only ever observed offering trips from Lyon to Grenoble and from Lyon to Paris, then we use characteristics of the Lyon to Paris trips when referring to the Lyon to Grenoble trip and using the term "trips other than the trip in question."

Using this approach, the instrument is the average price charged by the same driver on all trips other than the trip in question. For the drivers who only ever offer one trip, there are no such trips. We refer to these drivers as single-trip drivers and, for these drivers, the average-other-price instrument equals zero. This instrument is essentially a combination of two distinct characteristics of drivers. First, does the driver offer trips between a pair of cities that is different from the pair of cities in question? This factor

[10] We have explored including other variables to identify drivers uniquely, including whether a photo is shown and whether smoking is allowed in the car. Reassuringly, each approach classifies the overwhelming majority of drivers in the same way and all results that follow are robust to alternative driver classifications.

determines whether the instrument is positive. Second, if yes, did the driver set prices that were high or low, on average, on those other trips? This factor determines the continuous variation in the instrument if it is non-zero. In essence, our instrument is a combination of a dummy variable for whether the driver offers trips other than the trip in question and, if so, a continuous variable measuring average price on those trips.

We believe (and will provide evidence to support) that the instrument is very strong. The intuition behind the "average-other-price instrument" is that a combination of observed and unobserved characteristics of the driver affects the price she sets. However, the econometrician has access to all observed characteristics and thus the variation in price that is affected by the unobserved characteristics should be highly correlated across the driver's listings on the trip in question and her listings on trips other than the trip in question.

Further, we believe that the average-other-price instrument is plausibly exogenous because it reflects underlying factors about the driver that should not affect demand except through the price set on the listing in question. It is very useful that we have a large number of trips because constructing the average price the driver set for other listings on the same trip is likely to itself be endogenous; such an average-same-price instrument is problematic because potential riders might observe a given driver offering a given trip across multiple listings of the trip. By using the average price on trips other than the trip in question, we greatly reduce the possibility that riders have any sense of where the driver falls in the price distribution for other trips.

Using the average-other-price instrument, we conduct an instrumental variables analysis, where the first stage asks what factors affect the driver's price and the second stage asks what factors affect the quantity sold, controlling for the endogeneity of the price in its determination of quantity sold. Price is measured in integer euros and, as explained earlier, quantity sold is measured in two ways: the fraction of seats sold and a dummy variable that equals one if all seats sold.

Now return to Table 1, which shows summary statistics for the key variables in our data. The average price set by drivers is around 13 euros, with substantial variation (standard deviation of 9.4 euros). As discussed in the previous section, we have two measures of quantity sold: fraction sold measures the proportion of listed seats that sold (average of 62%), while all seats sold is a dummy variable (average of 53%). The instrumental variable used to handle the endogeneity of price is the average-other-price variable, which has an average of around 11 euros. Drivers offer 3.8 listings of their modal trip and 2.6 listings of trips other than their modal trip. Just over half of the drivers in our sample only ever offer one trip, implying that the average-other-price instrument equals zero for these drivers, which represent 21.1% of the total number of observations.

Figures 2 and 3 present the evolution of prices and the fraction of seats sold for all trips, on average, and for the single trip Paris-Lyon. Paris-Lyon is chosen because it is representative among the most commonly observed trips. Prices exhibit some volatility around a trend over time. Further, seats sold appear to positively covary with prices, reflecting underlying seasonality in both supply and demand. Finally, there is a noticeable decrease in prices and seats sold around the end of 2013 that recovers early in 2014. In Appendix A, we list the coefficient of variation by trip over the period, which measures price dispersion for a given trip. This dispersion measure ranges from 10.5% to 47.5% in these data.

[Insert Figures 2 and 3]

Econometric Specification

The econometric specification throughout uses fixed-effects panel-data regression, with trip fixed effects. For price, the model is a linear regression. For the fraction of seats sold and the all seats sold dummy variable, we again use linear regression. In both cases, the appropriate econometric specification is nonlinear: fractional logit in the case of fraction sold (which continuously varies between zero and one) and logit/probit in the case of all seats sold (which is a dummy variable). However, econometric models that handle endogeneity and allow for fixed effects are not available for either fractional logit or logit/probit. We could use a random-effect panel-data model but the orthogonality assumption on the unobserved effects imposed by the random-effects model does not hold in these data. Instead of ignoring endogeneity or ignoring unobserved trip-level effects, we use linear models throughout. Our approach is consistent with the approach advocated by Angrist and Pischke (2008). In all specifications, continuous explanatory variables are included in quadratic form, with the results shown as the marginal effect at the mean.

We now present the results from the first stage (price) and second stage (quantity sold).

4. Empirical Results

Table 2 presents the determinants of price, which serves as the first-stage of our instrumental variables regression analysis. Recall that our econometric specification is a panel-data regression with trip fixed effects. First, the average-other-price instrument (average price on trips other than the trip in question) is highly statistically significant.

Increasing the price on other trips by one euro is associated with a one cent higher price. The size of this effect suggests that the latent characteristics of the driver that introduce a correlation between prices on different trips are statistically meaningful but quantitatively small. This implies that trip-specific factors (e.g., the recommended price) are more important than driver-specific factors in pricing but driver-specific factors are sufficiently important to ensure the strength of the average-other-price instrument, as discussed now.

[Insert Table 2]

We discuss the validity of the average-other-price instrument using the first-stage F statistic that is shown in Table 2. The statistic equals 1712.5, which is very large and considerably above the rule-of-thumb that it exceed 10 in order to mitigate concerns about weak instruments (Angrist and Pischke 2008). Further, we measure the strength of both components of the variation in the instrument: first, the zero/non-zero variation in the instrument of whether the driver offers trips other than the trip in question and, second, the continuous variation in the instrument of the average price on those trips.

To measure the strength in the zero/non-zero variation, we rerun the first-stage weak instrument test with an instrument that equals one if the driver only ever offered a single trip. The first-stage F statistic in this case equals 179.2. To measure the strength in the continuous variation, we rerun the first-stage weak instrument test for only multiple-trip drivers to ask whether the continuous variation in the instrument is strongly associated with price. The statistic in this case equals 1670.2. These results suggest that most of the strength of the instrument comes from the continuous variation but that both sources of variation are sufficiently strong to support the use of the average-other-price instrument as a strong predictor of prices. In Appendix B, we discuss a set of robustness checks using two sets of alternative instruments. The results are very robust, as can be seen by comparing Table 3 to Tables B1 and B2.

To discuss the results, we discuss the determinants of price from Table 2 and of quantity sold from Table 3 together. In Table 3, results for the fraction sold measure are in Column (1) and for the all seats sold dummy in Column (2). The two sets of results are very similar, leading us to only discuss Column (1). Having a higher price is associated with fewer seats sold, where fraction sold decreases by around 8 percentage points for each one euro higher price, relative to a mean fraction sold of 62%.

[Insert Table 3]

We have two sets of main results from Tables 2 and 3: driver experience/reputation and driver demographics. After discussing these main results, we discuss the remaining findings.

Driver Experience/Reputation

The richness of our BlaBlaCar data allows us to control for driver reputation, which has been the focus on other studies of online markets, separately from driver experience. Reputation is measured in terms of its quantity (number of feedback ratings received) and quality (percentage of positive ratings relative to all ratings received). Holding reputation constant, driver status measures a driver's experience level on the BlaBlaCar platform.

The results suggest that more-experienced drivers set lower prices: drivers with the highest status (ambassador) set prices that are 44 cents lower than drivers with the lowest status (newcomer). This is a moderate effect size relative to a mean price of 13.4 euros; however, this effect represents one of the larger effects of any explanatory variable in Table 2. In contrast, driver reputation has a weak relationship with price. The quantity of a driver's reputation (more feedback) has an effect that is essentially zero, while the quality of a driver's reputation (better feedback) has a positive effect that is quite small: if a driver's reputation increases from 90% positive to 100% positive, for example, price is predicted to increase by two cents.

Turning to Table 3, drivers with more-experience sell more seats, controlling for price: ambassadors (highest status level) sell 5.7 percentage points more seats than newcomers (lowest status level). Drivers in the middle status levels (intermediate, experienced, and expert) sell around two percentage points more seats than inexperienced drivers but there is a discrete jump in demand for drivers at the highest experience level. Concerning reputation, more and better reputation is associated with higher quantity demanded: 10 additional ratings are associated with a 0.3 percentage point increase in fraction sold, while a 10 percentage point increase in reputation quality is associated with a two percentage point increase in fraction sold. Reputation matters more for demand than for price but its effects remain small nonetheless.

Overall, we conclude that drivers with better reputations (in terms of reputation quality) set higher prices, while drivers with better reputations (in terms of quantity and quality) sell more seats. However, the effects are much smaller than the effect of driver experience. More-experienced drivers set lower prices than less-experienced drivers, with a moderate effect size. Further, more-experienced drivers sell more seats, with a particularly strong effect associated with moving to the highest experience level.

Intuition from offline markets suggest that brand loyalty effects should allow more-established firms to charge higher prices. Evidence from eBay shows that seller with more experience are able to sell for higher prices (Cabral, 2016). In contrast, we interpret our finding as suggestive that new drivers on BlaBlaCar are using a different decision-making process when setting prices than that of experienced drivers. A plausible explanation is that new drivers are more attracted to BlaBlaCar by a profit motive, while experienced drivers have gained an appreciation for the nonpecuniary attributes of riders obtained through the platform (prosocial behavior).

Consistent with this interpretation, we find that the most-experienced drivers on BlaBlaCar (drivers with ambassador status) are more likely to display a photo (71.8% of ambassadors display their photo versus 33.1% of drivers with less experience) and play music (75.9% versus 52.0%). Further, the most-experienced drivers are much more likely to allow pets (13.5% versus 8.1%), despite the fact that pets are associated with lower prices (Table 2) and less demand (Table 3). Taken together, these results suggest that experienced drivers on BlaBlaCar are making choices that suggest the importance of nonpecuniary factors in their pricing decisions, rather than pecuniary motivations alone. We interpret this as evidence that prices and market outcomes on "sharing platforms" such as BlaBlaCar are determined differently than on other types of peer-to-peer markets such as eBay. Section 5 provides a discussion of potential mechanisms for the role of experience on pricing behavior.

Driver Demographics

Our next set of main results concern demographic characteristics of the driver: name origin, gender, and age. As discussed earlier, a driver's first/given name may signal her origin or ethnicity and we match names to predominant country of origin to classify drivers as having a French-sounding name (67% of drivers), an Arabic-sounding name (5%), or a name that is neither predominantly French nor Arabic (28%).

Drivers with a French name set prices that are essentially the same as the omitted group of all other names, while drivers with an Arabic name set prices that are around 19 cents lower. Controlling for price, drivers with a French name sell more seats, while drivers with an Arabic name sell fewer seats (fraction of seats sold increases by five percentage points and decreases by eight percentage points, respectively). We interpret these results in a similar vein as results from other online markets (e.g., Edelman and Luca, 2014), suggesting either discrimination or unobserved heterogeneity that is correlated with demographics. However, given our rich set of controls, we believe that there is limited scope for unobserved heterogeneity in explaining why we observe such differences because we control for essentially all of the characteristics that are observed by potential riders.

Next, we consider the interactive effects of these demographics with whether the driver uploaded a photo for riders to view. First note that drivers with a photo set prices that are no different from drivers without a photo, on average, while photos are associated with a very small increase in the fraction of seats sold. To ask whether demographics such as gender and name origin interact with the presence of a photo of the driver, we rerun the specification in Column (1) of Table 3 with a full set of interactions for gender by name origin by photo.[11]

The negative effect of having an Arabic-sounding name on seats sold is larger for drivers without a photo shown relative to drivers with a photo: demand is 7.9 percentage points lower for Arabic drivers than for non-Arabic drivers among drivers without a photo, but only 4.6 percentage points lower among drivers with a photo. Further, the negative effect is much larger for male drivers relative to female drivers: demand is 8.4 percentage points lower for Arabic males than for non-Arabic males, but only 2.9 percentage points lower for Arabic females than for non-Arabic females.

As a final check on these results for name origin, we rerun the specification in Column (1) of Table 3 with only those drivers whose car is of the highest comfort level (car class of four). One explanation of the demographic differences we observe is that they are explained by rider perceiving that a driver with a non-French name might drive a less luxurious car. In contrast to this hypothesis, among drivers whose car is of the highest comfort level, Arabic drivers also sell fewer seats than non-Arabic drivers and the effect size is similar to the main results (11.4 percentage point decrease in fraction sold). We conclude that the demographic results we find are driven by demographic preferences of riders for drivers who are female and have a non-Arabic name, where the latter effect is larger than the former. The disadvantage on BlaBlaCar of having an Arabic-sounding name is similar to the finding of a disadvantage on AirBnB faced by black hosts (Edelman and Luca, 2014).

[11] The full set of results from the robustness checks discussed in this section are available from the authors upon request.

Next, female drivers set prices that are 13 cents higher than male drivers, on average. Controlling for price, the fraction of seats sold is three percentage points higher for female drivers than for males. The higher demand for rides listed by female drivers may suggest that both female and male riders prefer a female driver but we do not have data on rider characteristics to test this hypothesis.

Finally, older drivers set higher prices and, controlling for price, sell no fewer seats, on average. Recall that age (along with the other continuous explanatory variables) is included in quadratic form, where the results shown in Table 2 are the marginal effect at the mean. To look for nonlinearities, Figures 4 and 5 present analyses of the effects of the driver age on price and quantity demanded (fraction of seats sold), respectively.[12]

[Insert Figures 4 and 5]

We ask whether the effect of age is linear, or whether incremental increases in the age of the driver are associated with different effects at different ages. For price, an additional year of age is associated with higher prices at all ages; the relationship shows minimal curvature, such that the slope is slightly diminishing in age. The magnitude of the effect is small but highly statistically significant, as can be seen by the 95% confidence intervals. In contrast to the linear effect on price, there is a nonlinear effect of age on quantity demanded: among younger drivers, an additional year of age is associated with more sales, while, among older drivers, an additional year of age is associated with fewer sales. The effect of age changes from positive to negative around 34 years of age.

[12] These results are generated using Stata's margins command and show the marginal effect of age on each outcome, along with 95% confidence intervals.

Putting the results on price and quantity demanded together, we see that drivers in their 30s fare better than drivers in their 20s, in terms of quantity demanded, despite the fact that drivers in their 30s set higher prices. Once drivers reach around age 40, the effect of further increases in age are associated with lower demand, yet these older drivers continue to set higher prices. This pattern is consistent with homophily, that is, younger drivers wanting to travel with young riders, perhaps because they share common interests. As younger riders are likely to be more price sensitive, young drivers may set lower prices to attract their peers. Similarly, older drivers are probably more selective in whom they ride with and high prices serve as a screening device.

Other Results

Beyond these main results, several other interesting patterns emerge. Class measures the car's comfort level, where zero represents no indication of the class, relative to values between one (basic comfort) to four (luxurious). The results suggest that, for drivers who do not disclose a class, prices are set as if the car is of average comfort (similar to a class of three). However, controlling for price, undisclosed quality cars are associated with fewer seats sold, where these cars sell one percentage point fewer seats than cars of the lowest disclosed comfort level. This is consistent with the unraveling result economists often predict under voluntary quality disclosure.

Concerning driver preferences, the following are associated with higher prices: not playing music, not allowing pets, allowing smoking, offering a roundtrip, and requiring manual confirmation.[13] The following are associated with more sales: playing music, not allowing pets, not allowing smoking, offering a roundtrip, and not requiring manual

[13] Manual confirmation requires a potential rider to request a ride from the driver, which must then be confirmed. 12% of drivers require manual confirmations. Other online markets have a similar option, including AirBnB.com, which allows hosts to require potential renters to request to book, rather than instantly book.

confirmation. The effect of manual confirmation is very large, suggesting a strong preference of riders not to be required to request confirmation of the ride.

Table 4 presents the final set of covariates, displaying the effects of departure time/day characteristics on price and quantity. That is, Column (1) of Table 4 presents results from the same regression as Column (1) of Table 2, while Columns (2) and (3) of Table 4 present results from the same regressions as Columns (1) and (2) of Table 3, respectively. For departure day of week, days with higher prices are also those days with more quantities sold, showing a strong preference for departures on Saturday and Sunday (the omitted day). For departure time in six hour intervals, we find that prices are lower for trips that depart later in the day, especially after 6PM, while more seats are sold between 6AM and 6PM, relative to nighttime hour departures.

Finally, a time trend is included to control for patterns over our eight month sample period. We find that prices trend downward, as do quantities; however, the aggregate data in Figure 2 do not suggest a downward trend other than a dip around the end of the calendar year. These trends are interesting but, as our sample is less than one year, we cannot control for an overall trend separately from monthly seasonality. As a result, we do not draw much in the way of interpretation of any potential time trend.

[Insert Table 4]

Finally, we present a robustness check of our two main results, where we present two analyses: (1) for each trip, a nonparametric trend test of whether prices fall as driver experience increases and (2) for each trip, a t-test of whether drivers with a predominantly Arabic name have lower average sales probability than other drivers.

First, across all 40 trips, we find that there is a statistically significant decrease in prices as driver experience increases for 37 of 40 trips (exceptions with a statistically significant increase: Lens-Paris, Nice-Toulon, and Rouen-Paris). We notice nothing systematic about these three exceptions (e.g., Lens-Paris has a positive price-experience relationship while Lille-Paris has a negative price-experience relationship, but the two trips are similar in most regards[14] as shown in Appendix Table A1). Further, specific interpretations present a concern about a multiple-testing problem (i.e., retesting over multiple subpopulations may reveal differences simply by chance).

Second, across all 40 trips, we find that there is a statistically significantly lower average sale probability for Arabic drivers for 35 of 40 trips (exception with a statistically significantly higher probability: Amiens-Beauvais, exceptions with no statistically significant difference: Besancon-Dijon, Dijon-Besancon, Metz-Nancy, and Saint Etienne-Clermont). Again, our focus is on the striking robustness of the main result of digital discrimination for drivers with a predominantly Arabic name.

The next section presents an analysis of mechanisms that potentially explain our first main result that prices fall as drivers gain experience.

5. Testing Explanations Based on Learning and Selection

Two separate mechanisms could account for the negative relationship we document between a driver's experience and her price. A first explanation involves learning: drivers learn to lower their prices as they gain experience. A second explanation involves a selection effect: drivers who offer low prices are more likely to gain a lot of experience. To provide evidence on the relative importance of these two explanations,

[14] The distance between Lille and Lens is only 36 kilometers.

we first use within-driver price dynamics to test for a learning effect, while accounting for selection. Then, we test for the degree to which selection exists.

Table 5 presents 12 sets of regression results that use the same specification but different subpopulations. The analysis explores the determinants of price-setting behavior following the first-stage regression results from Table 2. But here the regression includes driver fixed effects, which controls for the difference in price levels across drivers and exploits variation in price changes for a given driver over time. The question of interest is whether a given driver changes her price as she gains more experience, which is different than the earlier results that exploited variation across drivers. Earlier, we found that more-experienced drivers set lower prices than less-experienced drivers. Here we ask whether a driver sets lower prices when she has more experience relative to the prices she set when she had less experience.

[Insert Table 5]

Column (1) of Table 5 presents the results for all drivers with at least three listings, which includes 85,367 individual drivers and 685,648 observations. The remaining columns repeat the same regression for subpopulations of Column (1); the notes to the table describes each column's population in detail but, briefly, Columns (2)-(4) split drivers by the car's comfort level, Columns (5)-(6) by driver age, Columns (7)-(8) by driver gender, Columns (9)-(10) by driver name origin, and Columns (11)-(12) by driver music preference. In the econometric specification, the included covariates are driver feedback quantity and quality, driver experience, departure characteristics, trip fixed effects, and driver fixed effects (i.e., all time-varying covariates plus fixed effects).

Table 5 confirms that drivers lower their prices as they gain experience, thus supporting a learning explanation for our earlier results. Specifically, controlling for changes in a driver's feedback, moving to a higher level of experience is associated with lower prices. The other results in Table 5 demonstrate that this result holds robustly in many different subpopulations of drivers. The two cases in which more experience is not associated with monotonically lower prices are for drivers who do not disclose their car's comfort level (Column (2)) and for drivers with a predominantly Arabic name (Column (10)). While we could speculate about potential explanations for these two cases, we prefer to avoid doing so because of a concern about a multiple-testing problem, as mentioned earlier. The main message of Table 5 is that within-driver price changes drive our earlier result that more-experienced drivers set lower prices.[15]

Our interpretation is that drivers learn to set lower prices as they gain experience, suggesting that drivers learn about profit maximization over time (e.g., learn that lower prices increase demand enough to raise profits) and about utility maximization over time (e.g., learn that they enjoy the platform and the socialization it provides). To conclude our analysis of a learning explanation versus selection, we test the degree to which drivers with a lot of experience appear to be a selected sample.

To do so, we split each driver's tenure on BlaBlaCar during our sample period into two halves (e.g., for a driver with nine listings, the first four constitute her first half and the last five constitute her second half). Then, we summarize the driver's first-half prices to categorize drivers into four quartiles based on their relative prices during the first half

[15] The results in Table 5 are robust to including only those drivers we observe moving from a status of one to a status of five (i.e., progressing from a Newcomer near the beginning of our sample period to an Ambassador). There are 23,594 such drivers (310,097 total listings). Results using only these drivers are available from the authors upon request. Relative to when they were of status level one, drivers lower their prices by -0.19, -0.29, -0.51, and -0.50 when achieving status levels two through five, respectively.

of their tenure. To deal with price differences across trips and other factors that affect prices, we use the earlier first-stage regression model of prices, calculate the residuals (i.e., the driver's price on each listing minus the predicted price from the regression model), then find the average residual price for each driver during the first half of her tenure. Drivers are classified into four first-half price categories based on the four quartiles of the average residual prices, where the lowest (highest) category includes prices that were meaningfully lower (higher) than predicted and the middle two categories includes prices that were slightly lower (higher) than predicted.

The results in Table 6 suggest that there is not a monotonic pattern in a driver's number of second-half listings as a function of her first-half prices: the highest-priced drivers (category four) have 3.93 listings during the second half of their tenure, on average, while the lowest-priced drivers (category one) have 3.97 second-half listings. This difference is not statistically significant, which is inconsistent with the selection explanation (which says that the most-experienced drivers are a selected sample of drivers who consistently set low prices). Further, the two intermediate categories of sellers, drivers who set prices slightly lower than predicted (category two) or slightly higher than predicted (category three), have the most second-half listings.[16]

[Insert Table 6]

[16] The results in Table 6 are robust to including only those drivers we observe moving from a status of one to a status of five and are available from the authors upon request. The pattern of second-half listings across first-half price categories is the same as the included results: the first through fourth categories have 6.07, 7.54, 7.28, and 6.14 listings, respectively.

These results suggest that selection is not driving our results. Along with the results in Table 5, we conclude that drivers learn to set lower prices as they gain more experience on the platform.

6. Discussion and Conclusion

Despite the increasing importance of carsharing, these platforms have received only limited economic analysis. Our paper studies the largest carsharing platform in the world to understand the functioning of these type of peer-to-peer markets. We are particularly interested in assessing how much of our understanding from the literature on other types of peer-to-peer markets such as eBay carries over to the "new sharing economy" such as BlaBlaCar. The peer-to-peer markets in the latter category allow both online and offline interactions between users. As a result, price-setting and demand behavior are likely to present novel insights relative to the large literature that studies these questions using data from electronic marketplaces such as eBay.

Our analysis focuses on the leading carsharing platform, BlaBlaCar, which allows sharing of an intercity trip by connecting drivers with empty seats to potential riders. The main advantage of using BlaBlaCar to study pricing and market outcomes is that prices are set by individual drivers, relative to a "recommended price" that is suggested by BlaBlaCar. In contrast, on other peer-to-peer markets in the transportation sector such as Uber and Lyft, price setting is centralized and thus any driver offering a given trip at a given moment has the same price. Our focus is on experience and the potential for social motives in pricing setting, thus decentralized pricing is important to understand strategic behavior.

In an econometric model that explicitly accounts for price endogeneity, we find that more-experienced drivers set lower prices and, controlling for price, sell more seats. Further, we find that driver demographics matter in interesting ways: our quantitatively strongest demographic predictor of demand is whether the driver has an Arabic name, which robustly reduces the driver's quantity demanded.

The rich nature of our BlaBlaCar data allows us to present a detailed analysis of market outcomes in an important type of peer-to-peer market. However, as usual with data from online markets, there are some features of the data that limit the questions we can ask. First, we do not observe information about the riders who are buying seats. Thus, we cannot measure the degree of homophily or social links between drivers and their passengers. Second, we have only a binary scale for ratings (positive or negative). Since the time of our data collection, BlaBlaCar has adopted a five-level reputation measure, which would be useful to verify our findings on the effects of reputation on price and quantity demanded. Moreover, as the platform has matured, it will be interesting to analyze how the role of driver experience and demographics has evolved.

Our study represents a first step toward an understanding of pricing behavior and market outcomes on "sharing platforms" such as BlaBlaCar relative to the large literature on other types of peer-to-peer markets such as eBay.

References

Angrist, J. D., Pischke, J. S. (2008). Mostly Harmless Econometrics: An Empiricist's Companion. Princeton University Press.

Benabou, R., Tirole, J. (2006). Incentives and Prosocial Behavior. The American Economic Review, 96(5), 1652-1678.

Bolton, G., Greiner, B., Ockenfels, A., (2013). Engineering Trust: Reciprocity in the Production of Reputation Information, Management Science, 59, 2, 265-285.

Cabral, L., Hortaçsu, A. (2010). The Dynamics of Seller Reputation: Evidence from eBay, The Journal of Industrial Economics, 58: 54–78.

Cabral, L. (2012). Reputation on the Internet, in Martin Peitz and Joel Waldfogel (Eds), The Oxford Handbook of the Digital Economy, Chapter 13.

Dellarocas C., (2003). The Digitization of Word-of-Mouth: Promise and Challenges of Online Reputation Mechanisms, Management Science 49, p. 1407-1424.

Edelman, M., Luca, M. (2014). Digital Discrimination: The Case of airbnb.com, Harvard Business School, Working Paper 14-054.

Edelman B. G., Geradin D. (2015). Efficiencies and Regulatory Shortcuts: How should we regulate companies like AirBnB and Uber? Harvard Business School Working Paper 16-026.

Einav L., Farronato C., Levin J. (2016). Peer-to-Peer Markets, Annual Review of Economics, 8, September, forthcoming.

Evans, D. S., Schmalensee, R., Noel, M. D., Chang, H. H., & Garcia-Swartz, D. D. (2011). Platform economics: Essays on multi-sided businesses. Competition Policy International.

Fradkin, Andrey, Elena Grewal, David Holtz, Matthew Pearson. 2014. Reporting Bias and Reciprocity in Online Reviews: Evidence from Field Experiments on AirBnB. Working paper. Available at http://andreyfradkin.com/assets/long_paper.pdf

Gneezy, U., Meier, S., Rey-Biel, P. (2011). When and why incentives (don't) work to modify behavior. The Journal of Economic Perspectives, 25(4), 191-209.

Horton, J. J., Zeckhauser, R. J. (2016). Owning, Using and Renting: Some Simple Economics of the" Sharing Economy. National Bureau of Economic Research Working Paper No. 22029.

Jin, G., Kato A. (2006). Price, Quality and Reputation: Evidence from an Online Field Experiment. Rand Journal of Economics 37, p. 983-1004.

Jolivet, G., Jullien, B., Postel-Vinay, F. (2016). Reputation and Prices on the e-market: Evidence from a Major French Platform. International Journal of Industrial Organization 45, 59-75.

Melnik, M. I. and Alm, J. (2002), Does a Seller's eCommerce Reputation Matter? Evidence from eBay Auctions. The Journal of Industrial Economics, 50: 337–349.

Resnick, P., Zeckhauser, R. (2002). Trust Among Strangers in Internet Transactions: Empirical Analysis of eBay's Reputation System, M. Baye (Ed), Advances in Applied Microeconomics (Volume 11), Elsevier Science, Amsterdam, p. 667-719.

Resnick, P., Zeckhauser, R., Swanson, J., Lockwood, K. (2006). The Value of Reputation on eBay: a controlled experiment, Experimental Economics 9, p. 79-101.

Rochet, J.-C., Tirole, J. (2006). Two-Sided Markets: A Progress Report. Rand Journal of Economics 37: 645-667.

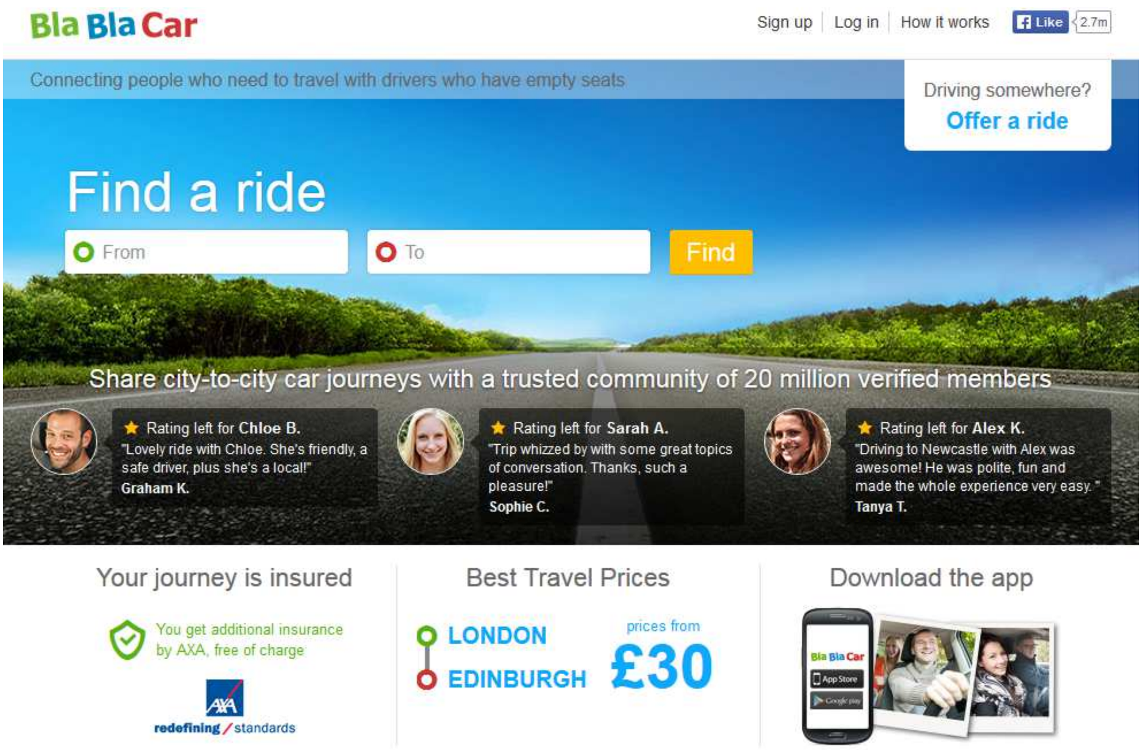
Schor J., Fitzmaurice, C. (2015) Collaborating and Connecting: The emergence of the sharing economy, Handbook on Research on Sustainable Consumption, eds., Lucia Reisch and John Thogersen, (Cheltenham, UK: Edward Elgar), 2015.

Sundararajan, A. (2016). *The Sharing Economy: The End of Employment and the Rise of Crowd-Based Capitalism.* MIT Press.

Zervas, G., Proserpio D., Byers J. W. (2014). The rise of the sharing economy: Estimating the impact of AirBnB on the hotel industry. Working paper. Available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2366898.

Zervas, G., Proserpio D., Byers J. W. (2015). A First Look at Online Reputation on AirBnB, Where Every Stay is Above Average. Available at http://ssrn.com/abstract=2554500

Figure 1: Screenshots of BlaBlaCar Website



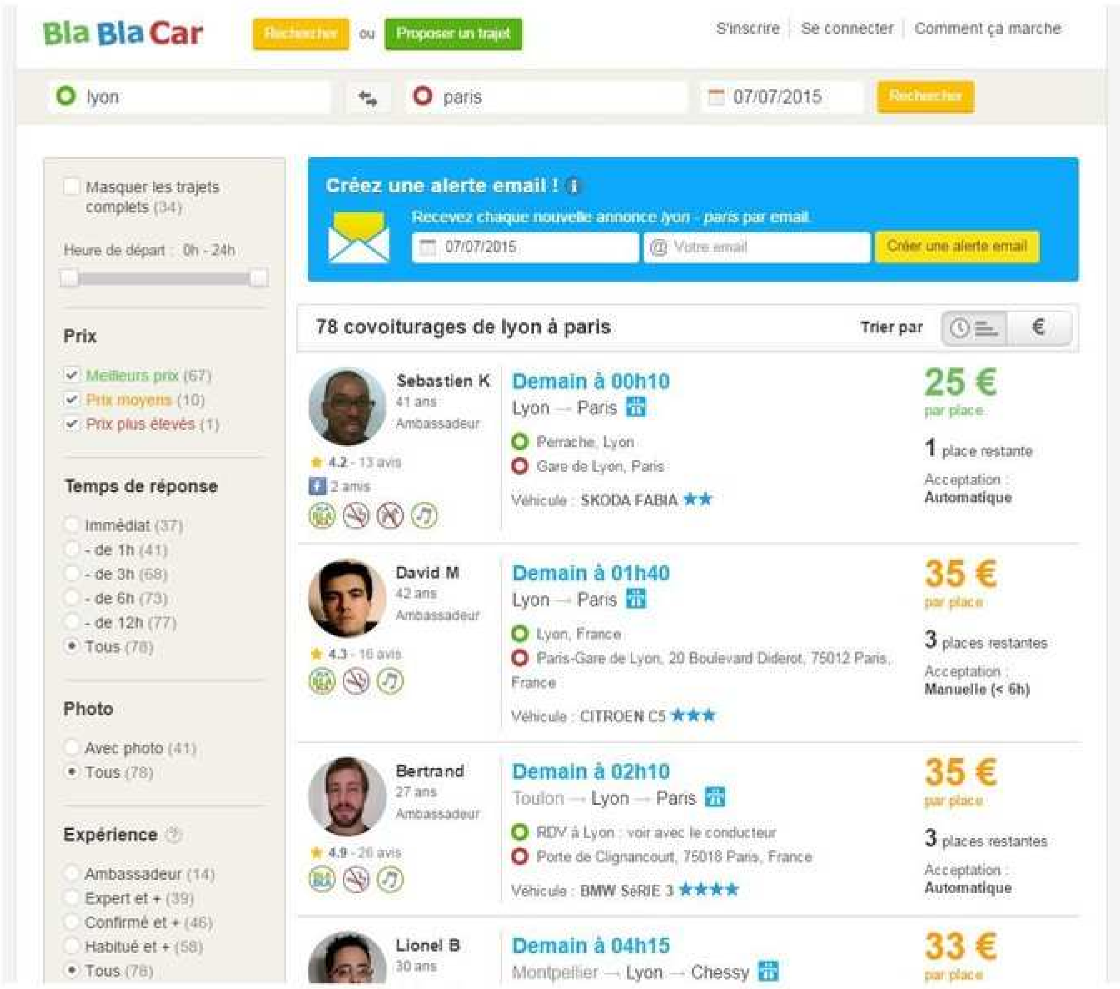


Figure 2: Average Fraction of Seats Sold and Prices Over Time for All Trips

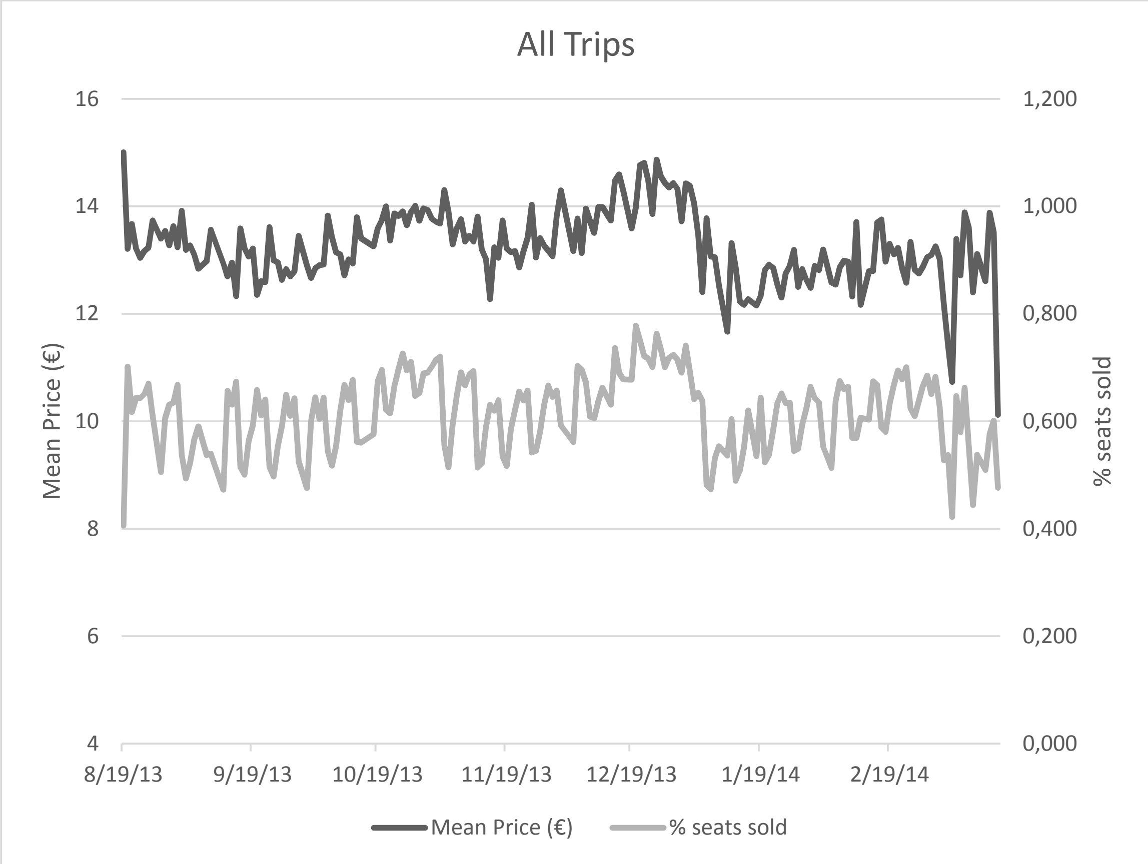


Figure 3: Average Fraction of Seats Sold and Prices Over Time for Paris-Lyon

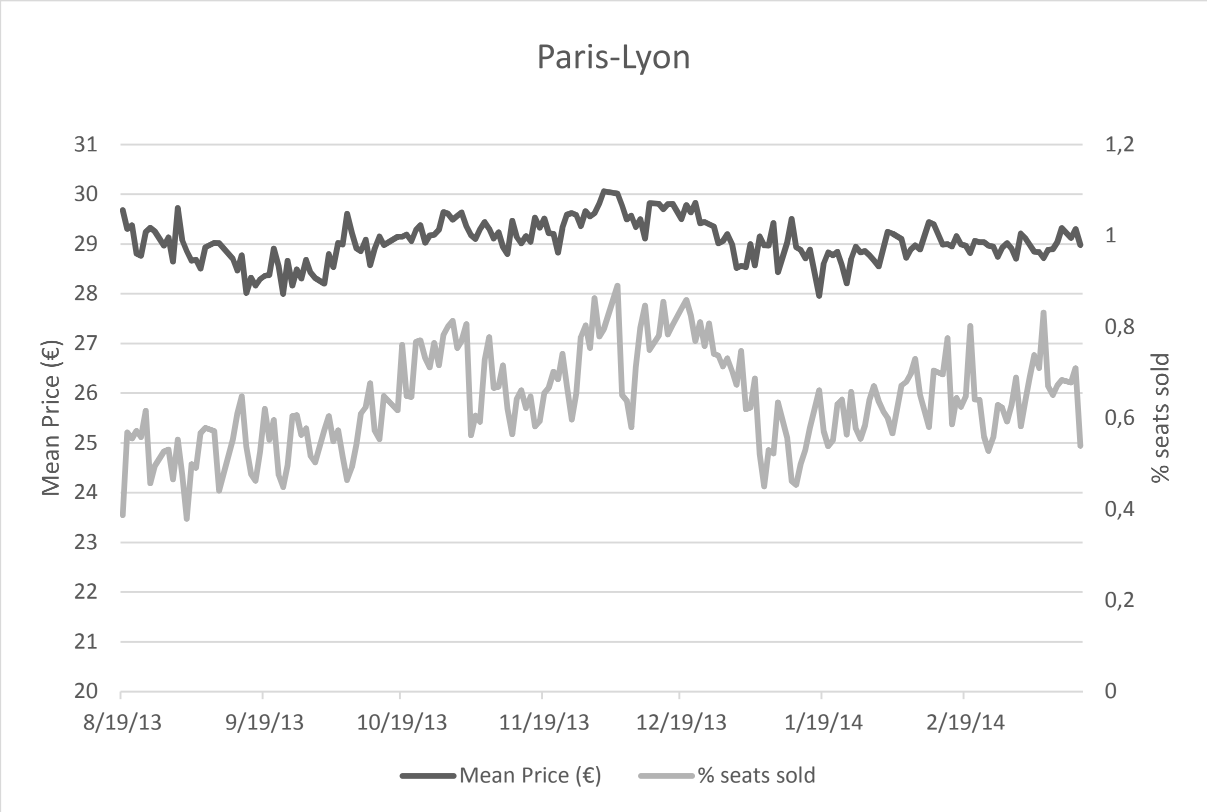


Figure 4: Nonlinearities in the Effect of Age on Price

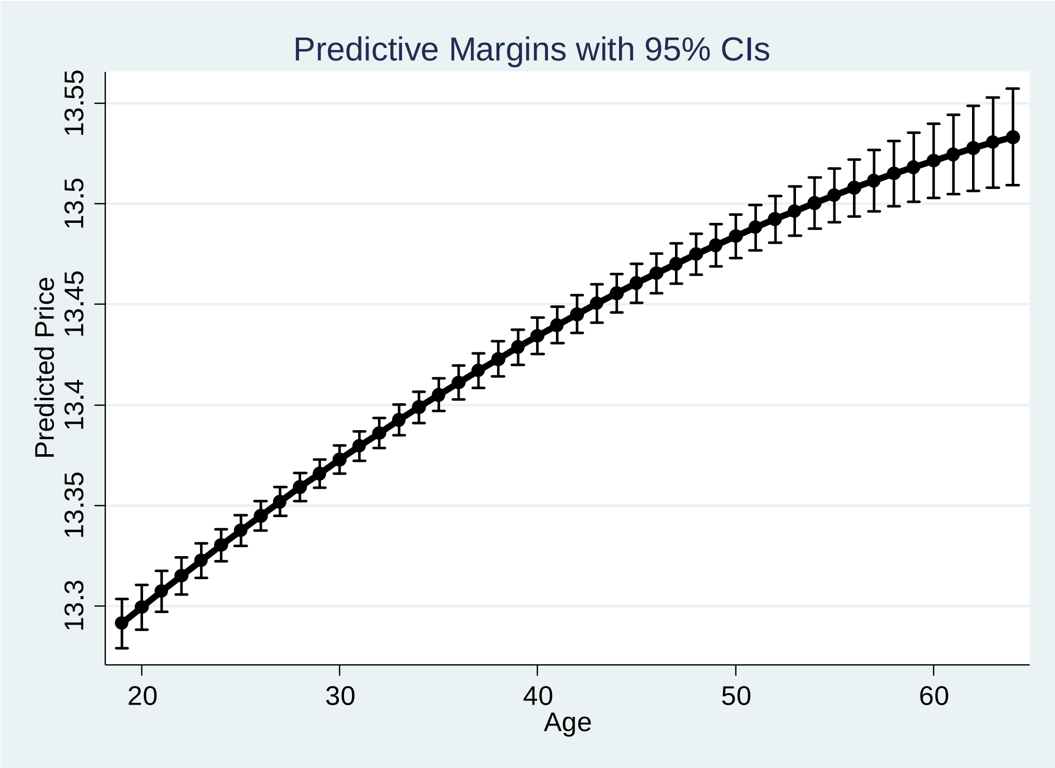


Figure 5: Nonlinearities in the Effect of Age on Pr(Sale)

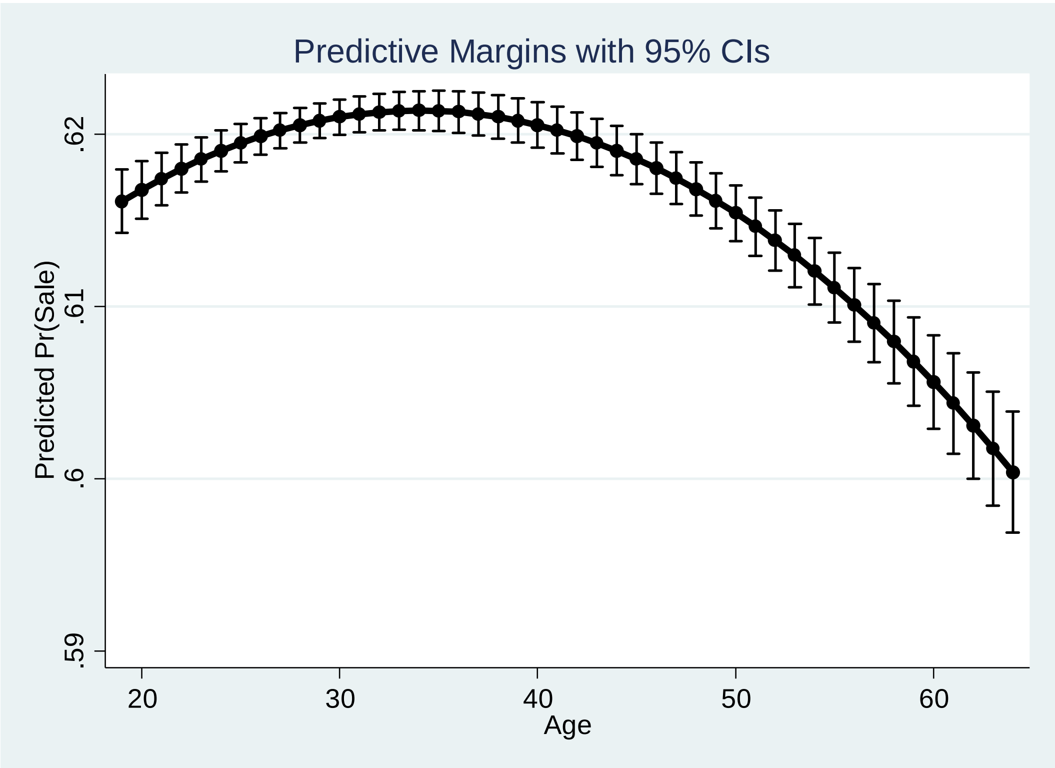


Table 1: Summary Statistics

Panel A: Unit of Observation = Trip	
Price	13.386
	(9.421)
Fraction Sold	0.618
	(0.435)
All Seats Sold	0.529
	(0.499)
Avg. Price, Other Trips	11.149
	(7.867)
N	948789
Panel B: Unit of Observation = Driver	
Number of Same Trips	3.787
	(6.944)
Number of Other Trips	2.565
	(6.760)
Single-Trip Driver	0.547
	(0.498)
Feedback Quantity	8.345
	(18.958)
Feedback Quality	99.037
	(6.080)
Driver Status	2.655
	(1.446)
Car Class	2.263
	(1.053)
Age	36.034
	(13.307)
Female	0.401
	(0.490)
French Name	0.673
	(0.469)
Arabic Name	0.047
	(0.212)
Photo Shown	0.388
	(0.487)
Plays Music	0.555
	(0.497)
Allows Pets	0.088
	(0.284)
Allows Smoking	0.067
	(0.250)
Roundtrip	0.259
	(0.438)
Manual Confirmation	0.116
	(0.320)
N	297582

Notes : Standard deviations are in parentheses.

Table 2: First-Stage Regression Results

	(1) Price
Avg. Price, Other Trips	0.009
	(0.000)***
Feedback Quantity	-0.000
	(0.000)
Feedback Quality	0.003
	(0.001)**
Driver Status=2	-0.149
	(0.009)***
Driver Status=3	-0.276
	(0.010)***
Driver Status=4	-0.396
	(0.010)***
Driver Status=5	-0.445
	(0.012)***
Car Class=1	-0.329
	(0.018)***
Car Class=2	-0.189
	(0.012)***
Car Class=3	0.017
	(0.012)
Car Class=4	0.314
	(0.016)***
Age	0.006
	(0.000)***
Female	0.130
	(0.007)***
French Name	0.014
	(0.009)
Arabic Name	-0.193
	(0.018)***
Photo Shown	-0.010
	(0.007)
Plays Music	-0.114
	(0.007)***
Allows Pets	-0.204
	(0.011)***
Allows Smoking	0.143
	(0.013)***
Roundtrip	0.168
	(0.007)***
Manual Confirmation	0.509
	(0.008)***
N	948789
First-Stage F Stat	1706.587

Notes: For this and subsequent tables, standard errors are in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: Main Regression Results

	(1) Fraction Sold	(2) All Seats Sold
Price	-0.081	-0.070
	(0.002)***	(0.003)***
Feedback Quantity	0.000	0.000
	(0.000)***	(0.000)***
Feedback Quality	0.002	0.002
	(0.000)***	(0.000)***
Driver Status=2	0.019	0.015
	(0.001)***	(0.001)***
Driver Status=3	0.015	0.012
	(0.002)***	(0.002)***
Driver Status=4	0.015	0.012
	(0.002)***	(0.002)***
Driver Status=5	0.057	0.064
	(0.002)***	(0.002)***
Car Class=1	0.012	0.016
	(0.003)***	(0.003)***
Car Class=2	0.032	0.036
	(0.002)***	(0.002)***
Car Class=3	0.030	0.037
	(0.002)***	(0.002)***
Car Class=4	0.039	0.042
	(0.002)***	(0.003)***
Age	0.000	0.000
	(0.000)	(0.000)***
Female	0.027	0.051
	(0.001)***	(0.001)***
French Name	0.061	0.026
	(0.001)***	(0.001)***
Arabic Name	-0.077	-0.090
	(0.003)***	(0.003)***
Photo Shown	0.005	0.006
	(0.001)***	(0.001)***
Plays Music	0.020	0.021
	(0.001)***	(0.001)***
Allows Pets	-0.017	-0.014
	(0.002)***	(0.002)***
Allows Smoking	-0.001	-0.006
	(0.002)	(0.002)***
Roundtrip	0.009	0.017
	(0.001)***	(0.001)***
Manual Confirmation	-0.429	-0.539
	(0.002)***	(0.002)***
N	948789	948789

Table 4: Regression Results of Departure Characteristics

	(1) Price	(2) Fraction Sold	(3) All Seats Sold
Departure=Monday	-0.128	-0.062	-0.065
	(0.011)***	(0.002)***	(0.002)***
Departure=Tuesday	-0.161	-0.110	-0.116
	(0.013)***	(0.002)***	(0.002)***
Departure=Wednesday	-0.110	-0.101	-0.111
	(0.013)***	(0.002)***	(0.002)***
Departure=Thursday	-0.009	-0.065	-0.071
	(0.012)	(0.002)***	(0.002)***
Departure=Friday	-0.054	-0.026	-0.028
	(0.009)***	(0.001)***	(0.001)***
Departure=Saturday	0.067	0.004	0.002
	(0.011)***	(0.002)***	(0.002)
Departure=[6AM,12PM]	-0.012	0.048	0.062
	(0.023)	(0.003)***	(0.003)***
Departure=[12PM,6PM]	-0.082	0.047	0.062
	(0.023)***	(0.003)***	(0.003)***
Departure=[6PM,12AM]	-0.232	0.004	0.015
	(0.024)***	(0.003)	(0.004)***
Departure Time Trend	-0.006	-0.002	-0.001
	(0.000)***	(0.000)***	(0.000)***
N	948789	948789	948789

Table 5: Regression Results of Within-Driver Price Changes

	(1)	(2)	(3)	(4)
Feedback Quantity	-0.001	-0.013	-0.007	-0.002
	(0.000)***	(0.003)***	(0.001)***	(0.000)***
Feedback Quality	-0.001	-0.008	0.002	0.001
	(0.002)	(0.007)	(0.003)	(0.002)
Driver Status=2	-0.167	-0.141	-0.139	-0.113
	(0.012)***	(0.050)***	(0.019)***	(0.018)***
Driver Status=3	-0.272	-0.257	-0.200	-0.244
	(0.012)***	(0.068)***	(0.020)***	(0.019)***
Driver Status=4	-0.407	-0.085	-0.332	-0.378
	(0.012)***	(0.081)	(0.023)***	(0.019)***
Driver Status=5	-0.435	-0.069	-0.357	-0.378
	(0.015)***	(0.158)	(0.034)***	(0.021)***
N	685648	60473	272316	352859
	(5)	(6)	(7)	(8)
Feedback Quantity	-0.001	-0.002	-0.001	-0.010
	(0.000)**	(0.000)***	(0.000)***	(0.002)***
Feedback Quality	-0.000	-0.002	-0.002	-0.000
	(0.002)	(0.002)	(0.002)	(0.003)
Driver Status=2	-0.199	-0.122	-0.195	-0.087
	(0.015)***	(0.018)***	(0.014)***	(0.021)***
Driver Status=3	-0.312	-0.212	-0.283	-0.199
	(0.016)***	(0.019)***	(0.015)***	(0.023)***
Driver Status=4	-0.436	-0.362	-0.437	-0.253
	(0.016)***	(0.019)***	(0.015)***	(0.027)***
Driver Status=5	-0.465	-0.392	-0.482	-0.184
	(0.020)***	(0.022)***	(0.017)***	(0.040)***
N	373512	312136	483118	202530
	(9)	(10)	(11)	(12)
Feedback Quantity	-0.002	-0.009	-0.009	-0.001
	(0.000)***	(0.003)***	(0.001)***	(0.000)**
Feedback Quality	-0.004	0.011	0.004	-0.000
	(0.002)**	(0.008)	(0.003)	(0.002)
Driver Status=2	-0.176	-0.220	-0.160	-0.106
	(0.012)***	(0.076)***	(0.018)***	(0.017)***
Driver Status=3	-0.279	-0.378	-0.243	-0.214
	(0.013)***	(0.086)***	(0.021)***	(0.017)***
Driver Status=4	-0.415	-0.256	-0.363	-0.346
	(0.013)***	(0.092)***	(0.024)***	(0.017)***
Driver Status=5	-0.442	0.199	-0.239	-0.369
	(0.016)***	(0.125)	(0.035)***	(0.019)***
N	592325	19725	275515	410133

Notes: These 12 sets of results present the within-driver price regression results for different subpopulations: Column (1) uses the entire sample of drivers with at least three listings, (2) such drivers with undisclosed quality cars, (3) drivers with cars rated one or two, (4) drivers with cars rated three or four, (5) drivers aged less than 30, (6) drivers aged 30 or older, (7) male drivers, (8) female drivers, (9) drivers with a predominantly French name, (10) drivers with a predominantly Arabic name, (11) drivers who do not play music on the trip, and (12) drivers who play music on the trip.

Table 6: Summary Statistics of Second-Half Listings Relative to First-Half Residual Prices

	(1)	(2)	(3)
	N	Price Range	Listings
Price Category=1	21342	<-0.752	3.967
			(0.031)
Price Category=2	21342	[-0.752,0.005]	4.722
			(0.031)
Price Category=3	21342	[0.005,0.695]	4.611
			(0.031)
Price Category=4	21341	>0.695	3.932
			(0.031)

Notes: These summary statistics include all drivers with at least three listings, showing price categories based on residual prices set during the first half of each driver's tenure during our sample period. Further, listings are the count of listings during the second half of each driver's tenure. Price categories are determined by quartiles of residual prices, which are the driver's price for each listing during the first half of her tenure minus the listing's predicted price from the earlier first-stage regression model.

Appendix A: Summary Statistics by Trip

Table A1: Summary Statistics by Trip

Trips	Number of observations	Distinct drivers	Average price	Coefficient of variation	Percentage Sale	Sold all	Distance (km)
Aix-Avignon	15407	10701	5.188	0.290	0.581	0.500	81
Amiens-Beauvais	4485	2472	4.264	0.282	0.546	0.472	66
Angers-Le Mans	23635	16248	6.184	0.223	0.552	0.460	97
Besancon-Dijon	3274	2241	5.893	0.357	0.444	0.342	93
Bordeaux-Nantes	37304	27029	21.169	0.111	0.619	0.521	353
Brest-Saint Brieuc	8049	5059	8.055	0.211	0.440	0.314	144
Caen-Rennes	23293	14371	10.987	0.158	0.499	0.378	185
Clermont-Lyon	29589	20859	12.351	0.161	0.555	0.446	166
Dijon-Besancon	3430	2294	5.876	0.337	0.434	0.330	93
Grenoble-Lyon	24507	15337	6.850	0.217	0.498	0.383	105
Le Havre-Caen	8984	5032	5.841	0.185	0.512	0.388	96
Le Mans-Tours	14716	10694	5.651	0.252	0.538	0.440	102
Lens-Paris	45467	27644	14.306	0.118	0.585	0.487	199
Lille-Paris	50244	29957	14.559	0.111	0.598	0.504	220
Limoges-Toulouse	20093	14779	18.015	0.169	0.543	0.437	291
Lyon-Grenoble	24544	15589	6.889	0.244	0.489	0.370	105
Lyon-Paris	49602	35312	29.155	0.117	0.661	0.573	466
Marseille-Nice	5283	3477	13.213	0.238	0.379	0.271	198
Metz-Nancy	32984	22394	3.544	0.279	0.923	0.912	60
Montpellier-Mars	15636	10843	10.911	0.196	0.506	0.405	169
Nancy-Strasbourg	4426	2960	9.206	0.250	0.453	0.339	156
Nantes-Bordeaux	37241	26965	21.136	0.108	0.609	0.510	353
Nantes-Rennes	53147	28492	5.868	0.237	0.501	0.386	113
Nice-Toulon	2086	1159	10.935	0.257	0.298	0.161	150
Nimes-Montpellier	91661	59904	3.071	0.266	0.962	0.957	58
Orleans-Paris	25468	19028	8.270	0.296	0.544	0.445	133
Paris-Brest	17447	13044	36.114	0.105	0.708	0.636	591
Paris-Caen	18308	11947	15.006	0.133	0.520	0.411	234

Paris-Lyon	48023	34552	29.117	0.126	0.654	0.566	466
Paris-Marseille	5471	4798	47.387	0.206	0.604	0.512	774
Pau-Bordeaux	12665	8361	13.927	0.187	0.460	0.336	218
Perpignan-Narbonne	24071	16778	4.197	0.300	0.744	0.698	66
Reims-Troyes	9472	6516	8.086	0.241	0.460	0.338	127
Rennes-Brest	11546	7170	12.763	0.198	0.458	0.325	243
Rouen-Paris	14182	8205	8.772	0.171	0.502	0.392	136
Saint Etienne-Clermont	36819	22047	22.191	0.475	0.562	0.451	144
Strasbourg-Colmar	8780	6449	4.337	0.269	0.713	0.664	76
Toulon-Aix	11237	7268	5.488	0.325	0.518	0.414	84
Toulouse-Bordeaux	48769	30144	15.114	0.134	0.535	0.413	245
Tours-Paris	27444	20069	15.441	0.184	0.624	0.536	240
Total	**948789**		**13.386**		**0.618**	**0.529**	

Appendix B: Robustness Checks Using Alternative Instruments

In this appendix, we introduce two strategies for constructing alternative instruments and demonstrate the robustness of the main results. First, we use the time series dimension of the data and look at drivers who offered listings of the trip in question during the week prior to the week in question. If the driver in question listed the trip in the prior week also, she is excluded from the construction of these instruments. Specifically, we use the one week lag of characteristics of other drivers on the same trip in the prior week. First, we construct the average feedback quality of other drivers on the same trip in the prior week. Second, we construct the average number of drivers with club status (with is akin to a driver of the highest status level) on the same trip in the prior week. In this strategy for constructing alternative instruments, we do not use feedback quantity or ambassador status because these instruments do not pass validity tests (i.e., there is evidence that they are weak).

The results are in Table B1. The first column reproduces Column (1) from Table 3 for comparison. The alternative instrument results are in the next three columns: Feedback Quality in Column (2), Driver's Club Status in Column (3), and both Feedback Quality and Club Status in Column (4).

[Insert Table B1]

The second strategy for constructing alternative instruments is to use a similar strategy used for the average-other-price instrument of exploiting the panel nature of the data. Here we use the average characteristics of other drivers than the driver in question's modal trip (other than the trip in question) during the same week. First, we construct the average feedback quantity of other drivers on the driver in question's modal trip in the same week. Second, we construct the average number of drivers with club status on the modal trip in the same week. In this strategy for constructing alternative instruments, we do not use feedback quality or ambassador status because these instruments do not pass validity tests (i.e., there is evidence that they are weak).

The results are in Table B1. The first column reproduces Column (1) from Table 3 for comparison. The alternative instrument results are in the next three columns: Feedback Quantity in Column (2), Driver's Club Status in Column (3), and both Feedback Quantity and Club Status in Column (4).

[Insert Table B2]

In both tables, the main results are highly robust. First, driver experience matters more than driver reputation but both positively affect sales, with the largest effect coming from the move to ambassador status (i.e., the highest level of experience). Second, driver demographics matter, with females and drivers with French names selling more seats and drivers with Arabic names selling fewer seats. The price elasticity itself is robustly statistically significantly negative but its magnitude varies depending on the instrument. Most effects in Tables B1 and B2 suggest price effects of six to eight percentage points decrease in the fraction of seats sold for a one euro higher price; however, the effect in Column (3) of Table B1 is much smaller. That said, our goal is not to obtain a precise estimate for price elasticity but instead to control for the endogeneity of price and, controlling for price, undercover the effects of driver characteristics (experience and demographics) on demand. In this regard, all of the results carry over.

Table B1: Robustness Checks Using First Set of Alternative Instruments

	(1)	(2)	(3)	(4)
Price	-0.081	-0.080	-0.014	-0.031
	(0.002)***	(0.010)***	(0.007)*	(0.006)***
Feedback Quantity	0.000	0.000	0.000	0.000
	(0.000)***	(0.000)***	(0.000)***	(0.000)***
Feedback Quality	0.002	0.002	0.002	0.002
	(0.000)***	(0.000)***	(0.000)***	(0.000)***
Driver Status=2	0.019	0.019	0.030	0.027
	(0.001)***	(0.002)***	(0.002)***	(0.001)***
Driver Status=3	0.015	0.015	0.034	0.029
	(0.002)***	(0.003)***	(0.002)***	(0.002)***
Driver Status=4	0.015	0.016	0.042	0.035
	(0.002)***	(0.004)***	(0.003)***	(0.003)***
Driver Status=5	0.057	0.057	0.087	0.079
	(0.002)***	(0.005)***	(0.003)***	(0.003)***
Car Class=1	0.012	0.013	0.035	0.029
	(0.003)***	(0.004)***	(0.003)***	(0.003)***
Car Class=2	0.032	0.032	0.045	0.041
	(0.002)***	(0.003)***	(0.002)***	(0.002)***
Car Class=3	0.030	0.030	0.029	0.029
	(0.002)***	(0.002)***	(0.001)***	(0.001)***
Car Class=4	0.039	0.039	0.018	0.023
	(0.002)***	(0.004)***	(0.003)***	(0.003)***
Age	0.000	0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)***	(0.000)***
Female	0.027	0.027	0.019	0.021
	(0.001)***	(0.002)***	(0.001)***	(0.001)***
French Name	0.061	0.061	0.062	0.061
	(0.001)***	(0.001)***	(0.001)***	(0.001)***
Arabic Name	-0.077	-0.076	-0.065	-0.068
	(0.003)***	(0.003)***	(0.003)***	(0.002)***
Photo Shown	0.005	0.005	0.005	0.005
	(0.001)***	(0.001)***	(0.001)***	(0.001)***
Plays Music	0.020	0.020	0.028	0.026
	(0.001)***	(0.002)***	(0.001)***	(0.001)***
Allows Pets	-0.017	-0.017	-0.003	-0.007
	(0.002)***	(0.003)***	(0.002)*	(0.002)***
Allows Smoking	-0.001	-0.002	-0.011	-0.009
	(0.002)	(0.002)	(0.002)***	(0.002)***
Roundtrip	0.009	0.009	-0.003	0.000
	(0.001)***	(0.002)***	(0.001)*	(0.001)
Manual	-0.429	-0.430	-0.464	-0.455
Confirmation	(0.002)***	(0.006)***	(0.004)***	(0.003)***
N	948789	947052	947052	947052
First-Stage F Stat	1706.587	92.955	148.625	112.244

Notes: These robustness checks present alternative instruments with fraction of seats sold as the dependent variable. Column (1) reproduces Column (1) from Table 3 for comparison. These alternative instruments are based on the one week lag of average characteristics of other drivers on the same trip (excluding the driver in question), as described in the paper: Feedback Quality in Column (2), Driver's Club Status (akin to a driver of the highest status level) in Column (3), and both Feedback Quality and Club Status in Column (4).

Table B2: Robustness Checks Using Second Set of Alternative Instruments

	(1)	(2)	(3)	(4)
Price	-0.081	-0.073	-0.040	-0.062
	(0.002)***	(0.019)***	(0.013)***	(0.011)***
Feedback Quantity	0.000	0.000	0.000	0.000
	(0.000)***	(0.000)***	(0.000)***	(0.000)***
Feedback Quality	0.002	0.002	0.002	0.002
	(0.000)***	(0.000)***	(0.000)***	(0.000)***
Driver Status=2	0.019	0.020	0.026	0.022
	(0.001)***	(0.003)***	(0.002)***	(0.002)***
Driver Status=3	0.015	0.017	0.026	0.020
	(0.002)***	(0.006)***	(0.004)***	(0.003)***
Driver Status=4	0.015	0.018	0.032	0.023
	(0.002)***	(0.008)**	(0.005)***	(0.005)***
Driver Status=5	0.057	0.061	0.075	0.066
	(0.002)***	(0.009)***	(0.006)***	(0.005)***
Car Class=1	0.012	0.015	0.026	0.019
	(0.003)***	(0.007)**	(0.005)***	(0.004)***
Car Class=2	0.032	0.033	0.040	0.035
	(0.002)***	(0.004)***	(0.003)***	(0.003)***
Car Class=3	0.030	0.030	0.029	0.029
	(0.002)***	(0.002)***	(0.001)***	(0.002)***
Car Class=4	0.039	0.037	0.027	0.033
	(0.002)***	(0.007)***	(0.004)***	(0.004)***
Age	0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)***	(0.000)
Female	0.027	0.026	0.022	0.025
	(0.001)***	(0.003)***	(0.002)***	(0.002)***
French Name	0.061	0.061	0.061	0.061
	(0.001)***	(0.001)***	(0.001)***	(0.001)***
Arabic Name	-0.077	-0.075	-0.069	-0.073
	(0.003)***	(0.004)***	(0.003)***	(0.003)***
Photo Shown	0.005	0.005	0.005	0.005
	(0.001)***	(0.001)***	(0.001)***	(0.001)***
Plays Music	0.020	0.021	0.025	0.022
	(0.001)***	(0.002)***	(0.002)***	(0.002)***
Allows Pets	-0.017	-0.016	-0.009	-0.013
	(0.002)***	(0.004)***	(0.003)***	(0.003)***
Allows Smoking	-0.001	-0.003	-0.007	-0.004
	(0.002)	(0.003)	(0.002)***	(0.002)*
Roundtrip	0.009	0.007	0.002	0.005
	(0.001)***	(0.004)**	(0.002)	(0.002)***
Manual	-0.429	-0.433	-0.450	-0.439
Confirmation	(0.002)***	(0.010)***	(0.007)***	(0.006)***
N	948789	948789	948789	948789
First-Stage F Stat	1706.587	25.587	49.903	38.265

Notes: These robustness checks present alternative instruments with fraction of seats sold as the dependent variable. Column (1) reproduces Column (1) from Table 3 for comparison. These alternative instruments are based on the average characteristics of other drivers on the driver's modal trip (excluding the driver in question), as described in the paper: Feedback Quantity in Column (2), Driver's Club Status (akin to a driver of the highest status level) in Column (3), and both Feedback Quantity and Club Status in Column (4).