Fighting discrimination with reputation: The case of online platforms.[*]

Xavier Lambin[†] Emil Palikot[‡]

October 18, 2022

Abstract

This paper shows that reputation systems can mitigate ethnic discrimination by enabling ethnic minority sellers to accrue high reputation quickly, leading buyers to update their beliefs. Using data from a ridesharing platform, we find that minority drivers with no reviews make 12% less revenue relative to similar nonminority drivers. This disparity gradually shrinks and almost disappears for experienced drivers. To understand the mechanism behind this process, we construct a model of career concerns' of discriminated sellers in the presence of a reputation system. The model's estimates show that minority drivers, who just entered the platform, face overly pessimistic beliefs about the quality of their service. To alter these beliefs, they exert high effort and offer low introductory prices, swiftly boosting their reputation. Counterfactual simulations reveal that the cost of incorrect prior beliefs is high and that the reputation system strictly benefits minority drivers.

JEL Classification: J15, L14, L91

[*]We are grateful to Charles Angelucci, Mark Armstrong, Maria Elena Bontempi, Jacques Crémer, Daniel Ershov, Daniel Garrett, Jonas Hjort, Alex Imas, Marc Ivaldi, Bruno Jullien, Yassine Lefouili, Andrea Mantovani, Sarit Markovich, Nicolas Pistolesi, Imke Reimers, Patrick Rey, Mathias Reynaert, Tobias Salz, Mark Schankerman, Timothy Simcoe, Steven Tadelis and Jean Tirole for their valuable comments at various stages of the paper. We also thank participants to seminar sessions at Telecom ParisTech, Columbia University, Northeastern University, Boston University, Toulouse School of Economics and the IIOC, AFSE, EARIE, Jornadas de Economia Industrial, WIPE conferences for useful comments and suggestions.

[†]ESSEC Business School and THEMA, xavier.lambin@essec.edu.

[‡]Ph.D candidate at Toulouse School of Economics, emil.palikot@gmail.com. Emil Palikot acknowledges support from the European Research Council under the Grant Agreement no. 340903.

1 Introduction

The online economy promised to eliminate offline frictions and facilitate collaboration among strangers. Reputation systems (reviews and ratings) provide a key mechanism for this: by aggregating information about past transactions, they discipline buyer and seller behavior and favor high quality types (e.g., Tadelis (2016)).[1] This should ensure the efficient functioning of online markets. Yet, there exists substantial evidence of severe discrimination online. On Airbnb, black hosts charge less than non-black hosts for equivalent rentals, and booking requests from black guests are less likely to be accepted (Edelman and Luca (2014); Edelman et al. (2017)). The goal of this paper is to investigate this apparent contradiction.

We collect data on a ridesharing platform that reconcile these seemingly incompatible facts. We find evidence of ethnic discrimination against minority drivers but also observe that reputation-building, thanks to passenger reviews, allows drivers to overcome *initial* discrimination. Estimating a model of career concerns, we show that the reputation system does indeed enable minority drivers to mitigate the handicap from which they initially suffer. However, building a reputation comes at a cost; as a result, the foregone payoffs stemming from the initial prejudice appear to be quantitatively important.

To perform this study, we have collected data on BlaBlaCar, a prominent French carpooling platform. BlaBlaCar is mostly used for inter-city trips with an average length of 400 km. Hence, the rides typically lead to several-hour-long interactions. Two features of the platform design are critical to our analysis. First, passengers can indeed discriminate. When searching for a ride, passengers see the profiles of all available drivers, which include their names, photos, and all the reviews from previous rides. Second, drivers set prices and collect reputation. Thus, they can act to influence demand. By exerting effort to obtain positive reviews, and setting low prices to collect these reviews at a faster pace, they can boost their reputation.

Our data show sizable differences across ethnic groups in terms of listing popularity (measured by the number of clicks they generate), the number of seats sold, and revenue generated by the listing. This disparity is robust to a rich set of driver-and listing-specific controls. The gap is concentrated in

[1] A key feature underlying the success of the "sharing economy" is the efficacy of reputation systems in building trust across social divides. See a talk by Joe Gebbia, a co-founder of Airbnb: https://www.youtube.com/watch? v=16cM-RFid9U, last accessed October 22, 2019. Furthermore, Frederic Mazzella, BlaBlaCar CEO, claims that the company's reputation system creates "a sense of trust almost comparable to the level of trust in friends" (Mazzella and Sundararajan (2016)).

the beginning of drivers' careers and shrinks as they receive reviews. Ethnic minority drivers with fewer than five reviews earn twelve percent less revenue than do nonminority entrants. This difference declines to seven percent for drivers with more than five and fewer than fifteen reviews and is statistically insignificant for users with more than forty reviews. Third, we show that the change of sample composition due to the exit of underperforming minority drivers is not the mechanism behind our results. Fourth, the analysis of the within driver variations in prices and grades reveals that drivers set lower prices and receive higher grades when they are new on the platform. Both effects are stronger for minority than nonminority drivers.

To highlight the causal link between new reviews and improvements in the economic performance of minority drivers, we exploit a natural experiment consisting of demand shocks. We carry out a difference-in-differences analysis where the treated group used the platform during an event of extraordinarily high demand caused by a railway strike, while the control group used the platform on a regular (non-strike) day. The treatment is an exogenous increase in the number of reviews available on profiles of drivers that happened to be driving on a strike day. We find that the minority drivers in the treated group achieved substantially higher revenue after the treatment than did the minority drivers in the control group.

Minority drivers have a strong incentive to build a reputation. To study how they respond to this incentive by *investing* in reputation, and to evaluate the costs of the initial prejudice, we propose a model of career concerns. Our model builds on Holmström (1999); drivers, characterized by intrinsic types (initially incompletely known) and marginal costs, set prices and exert efforts to maximize life long consumption. Passengers observe a set of available drivers and choose the one that maximizes their expected utility. They have prior beliefs about the distribution of drivers' types, which are population-specific and might be incorrect. After a ride, the passenger reports truthfully the quality of service; the report is used in successive periods to form posterior beliefs about the driver's type. The quality of service is a function of the driver's type, the amount of effort she puts in, and a random shock. Passengers observe and report the overall quality, not the individual components.

Drivers' pricing and effort decisions exhibit static-dynamic tradeoffs: they can decide to offer discounts and exert costly effort to build a high reputation quickly. The incentive to invest in reputation is strong when passengers value reputation highly, and a grade has a substantial impact on posterior beliefs; the more randomness the reviews exhibit, the lower are the efforts. Furthermore, there are decreasing returns to investing in reputation because each subsequent grade has a smaller impact on posterior beliefs. As a result, both efforts and discounts tend to zero over time.

In a market defined as a day and route combination, we observe all available drivers, their characteristics, prices, and the number of sold seats. We also know how many times each listing was viewed by potential passengers, which gives us a precise measure of the number of passengers looking for a ride and allows us to model passengers' choice problem. Each passenger chooses a driver that maximizes her expected utility from a set of available drivers and the outside option. We estimate the parameters of demand by maximizing a loglikelihood function. The crucial assumption allowing us to identify the parameters of the supply is that after a certain number of reviews, enough information is available on drivers' profiles so that in the subsequent periods they do not exert effort or offer discounts.[2] We identify drivers' types and their marginal costs from grades and prices observed after the reputation building stage.

We use market outcomes to back out beliefs about the quality of service. We show that the market expects a minority driver with no reviews to be of quality 4.17 (i.e., 8th percentile of the distribution of grades) on a scale of 1 to 5 despite grading them after the trip 4.62 (48th percentile) on average. The disparity between the expected and given grades is the consequence of incorrect prior beliefs.

Prior beliefs influence incentives to invest in reputation. An additional review leads, on average, to an improvement in posterior beliefs about the quality of service of minority drivers. Consequently, minority drivers offer low introductory prices that increase the chance of selling a seat and being reviewed. The optimal prices that contain the component of investing in reputation are over eight percent lower than the price that would maximize current pay-off (the discount offered by nonminority entrants at the reputation building stage is four percent). The incentive to exert effort depends on future profits and the amount of uncertainty about the driver's type. Minority drivers initially have lower expected future profits, but there is higher uncertainty about their types. Considering both effects, we find that they have higher incentives to exert effort than nonminority drivers.

Establishing a reputation is costly as minority drivers have to go through an initial period of low outcomes and additionally need to *invest* in reputation building. In a counterfactual, we assume that passengers have correct prior beliefs about the quality of service offered by minority drivers. We can quantify the cost of the incorrect priors and resulting discrimination by comparing the counterfactual profits to the baseline scenario: we show that the average pay-off of minority drivers over the first fifteen rides is nineteen percent higher in the counterfactual case.

[2]The model shows that efforts and discounts tend to zero as drivers collect reviews. The within driver variation in prices and grades exhibit patterns consistent with investing in reputation until approximately the tenth review.

In a second counterfactual, we study what happens when the initial disparity between minority and nonminority drivers does not fade away. In this scenario, passengers always consider minority drivers to be of a lower quality.[3] As a result, minority drivers' incentives to invest in reputation vanish, they increase introductory prices and exert much less effort. Their average pay-off throughout the first fifteen rides is eight percent lower than the baseline.

Finally, we analyze the effects of the introduction of ethnicity-blind profiles, as proposed by Edelman et al. (2017). In this experiment, passengers are ex-ante uncertain whether a driver is from a minority or not. When passengers cannot establish the ethnicity of a driver based on the profile, there is no discrimination at the booking stage, which influences the prices and efforts of both minority and nonminority drivers. Minority drivers increase their prices and offer a better quality of service. Their profits increase substantially, nonminority drivers' profits are reduced.

Relation to literature: This paper relates to several strands of economic literature. First, the differences in economic outcomes across ethnic groups have been studied for a long time, see, e.g., Kuznets (1955); Alesina et al. (2016) show the extent of ethnic inequality worldwide. The negative impact of ethnic discrimination on economic outcomes is well documented: Banerjee and Munshi (2004) quantify the aggregate loss due to discriminatory investment decisions, and Hjort (2014) shows high economic costs of ethnic preferences in team production.[4] Discrimination against ethnic minorities in digital markets has been mostly studied in the context of short-term house rentals.[5] In the case of ridesharing, Farajallah et al. (2019) show that ethnic minority drivers set lower prices than nonminority drivers.[6] We contribute to this literature by documenting a gap in revenues and economic profits. Most importantly, we develop a model of belief formation and updating, which allows us to estimate incorrect prior beliefs and understand their impact on the economic outcomes of minority drivers. We also show that these beliefs are updated with reviews.[7] Furthermore, the analysis of counterfactuals

[3] The expected quality of service of individual minority drivers suggested by their average grades is always decreased by the size of the initial gap.

[4] The economic theory of discrimination generally follows two approaches. Taste-based discrimination, formalized by Becker (1971), attributes discrimination to preference against interacting with some economic agents. While, the theory of statistical discrimination, due to Phelps (1972) and Arrow (1973), explains discrimination in terms of differences in the expected quality across groups; when an individual agent's quality is not observed, the expectation of it is formed based on the observed minority status. The distinction between statistical discrimination with correct and incorrect priors has recently been discussed by Bohren et al. (2019a). Bohren et al. (2019b) formalizes the theory of dynamic discrimination.

[5] See: Edelman and Luca (2014), Edelman et al. (2017), Laouenan and Rathelot (2017), and Kakar et al. (2018).

[6] The majority of empirical work in this domain identifies a disparity in prices between minority and nonminority sellers. However, a difference in prices is not necessarily due to discrimination; we show that part of it can be explained by seller heterogeneity in unobserved characteristics, for example, marginal costs.

[7] The importance of information to minority groups is shown in experimental settings by Bartoš et al. (2016) and Cui et al. (2019). Agrawal et al. (2016) provide evidence that information benefits employees from less developed countries. The

allows us to quantify the cost of incorrect beliefs.

Second, our structural model builds on the literature on dynamic moral hazard. We generalize the seminal model of Holmström (1999) by introducing incorrect beliefs, competition between drivers, and pricing as an additional strategic tool.[8] The structural estimation of a career concerns model using data from a reputation system is our contribution to this literature. The estimation results allow us to study drivers' reactions to discrimination. Coate and Loury (1993) and Glover et al. (2017) argue that discrimination can be a self-fulfilling prophecy. We show that conditioned on entering the market, minority drivers facing statistical discrimination with an erroneous prior exert effort and set low introductory prices to improve their future outcomes.

Third, Ge et al. (2016) show that the magnitude of discrimination depends on how early, in the booking process, the information on ethnicity becomes available. Thus, the extent of discrimination varies with the design of a marketplace. Edelman et al. (2017) discuss various policy proposals aimed at mitigating discrimination online; such policy interventions spur reactions by all market participants. Our structural model allows us to generate counterfactuals and evaluate the welfare effects of various market designs.

The rest of this paper is organized as follows: section 2 introduces some important features of BlaBlaCar and the data collection process. Section **??** provides reduced-form results. We document the output gap between minority and nonminority drivers, the analysis of which is followed by a study of the effect of reputation building and a comparison of exit patterns. Next, we perform a difference-in-differences analysis exploiting a natural experiment. In section 5, we introduce a model of passenger choice and drivers' career concerns. Next, in section 6 we discuss identification assumptions and the estimation procedure. Section 7 presents the estimation results. Section 8 describes counterfactual experiments. Finally, we conclude the paper in section 9.

2 Empirical context and data collection

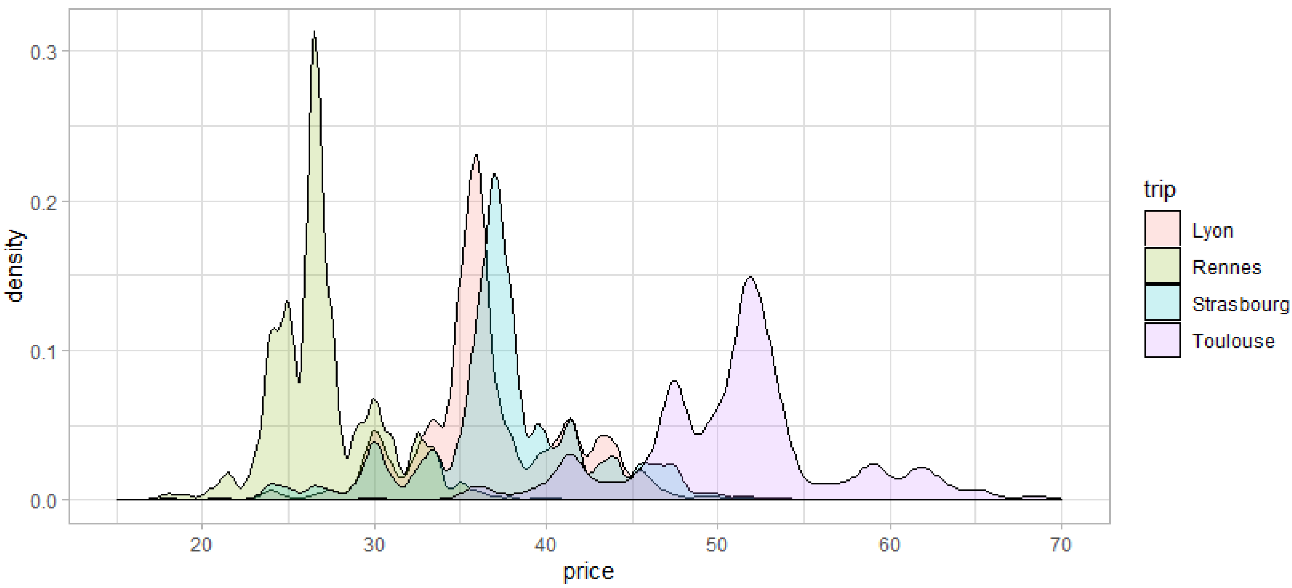
BlaBlaCar is an online marketplace for ridesharing that was established in 2006 in France and today operates in 22 countries, mostly in Europe, but also Mexico, India, and Brazil. The platform has over 100 million active users.[9] BlaBlaCar is particularly popular in France, where 1.5 million passengers use

additional benefit of acquiring information about new workers is explored by Pallais (2014). Sociological research has also studied the potential of reputation systems to offset trust judgments, see, e.g., Abrahao et al. (2017); Tjaden et al. (2018); Carol et al. (2019).

[8] Employer learning has been captured before by Chiappori et al. (1999) and Altonji and Pierret (2001).

[9] https://blog.blablacar.com/newsroom/news-list/blablacar-reaches-100-million-members-for-its-15th-anniversary

Figure 1: Price dispersion on BlaBlaCar



Note: Distribution of prices in euros on routes Paris to/from Lyon, Rennes, Strasbourg, Toulouse.

it every month. There are several essential differences between BlaBlaCar and ride-hailing services, such as Uber or Lyft, we discuss them in this section.

Participation in BlaBlaCar is restricted to nonprofessional drivers; this is ensured by imposing limits on the number of seats and listings drivers can offer.[10] Typically, drivers travel on a given route for personal reasons and use the platform to cover some of the costs. BlaBlaCar is particularly popular on long routes between major cities. In our dataset, the average trip is 400 km long. Thus, a decision to travel with someone implies interacting for several hours.

Another key feature of BlaBlaCar is that drivers' set their prices. BlaBlaCar offers a suggestion that depends only on the distance and amounts to 0.062 EUR per km. Drivers typically deviate from the suggestion.[11] Figure 1 shows the distributions of prices on several popular routes. There is a significant degree of price dispersion within routes.

Before booking a ride, a potential passenger sees a list of all drivers available on a given route. By default, drivers are ranked by departure time. Some basic information is displayed at this stage: the driver's photo, name, average rating, a few details about the ride, and the price. To obtain more information, and in particular, to see the history of reviews, a prospective passenger needs to click on and visit the profile of the driver.[12] The passenger chooses the listing that she finds the most attractive

[10] In 2019, after our sampling period, BlaBlaCar introduced BlaBlaBus, a professional bus service.

[11] The price is capped at 0.082 EUR per km, but this cap is very rarely binding.

[12] Examples of profiles and listing pages are provided in Appendix A.

and sends a booking request. Approximately half of the drivers choose the automatic acceptance feature while posting a ride; others reserve the option to reject requests. Finally, payment is made upfront via the BlaBlaCar online system. BlaBlaCar fees (see Appendix A) are deducted from the price paid by the passenger.

BlaBlaCar sends multiple reminders to encourage the passengers and the driver to leave reviews. A review consists of a textual comment and a grade from 1 to 5. We have collected both the written comments and grades. We carried out a sentiment analysis of the written comments; this exercise reveals that there is a high correlation between the sentiment expressed with a written review and the associated grade. We document this in Appendix D. Given this high correlation, we decided to focus only on grades. From now on, we will use *review*, *rating* and *grade* interchangeably while referring to a grade on the scale of 1 to 5.[13]

Reviews on sharing economy platforms are frequently skewed to the right (disproportionately positive). If a vast majority of reviews assign the highest possible grade, the reputation system loses its informativeness (Zervas et al. (2015) studies the implications of this). On the BlaBlaCar platform, we also see that the highest possible grade of 5 is the most popular. However, there are still enough reviews with lower grades to make the grading system meaningful. The mean grade per driver in our dataset is 4.6.

Data collection: We have collected our dataset using a web crawler on the website www.blablacar.fr, from 1.07.2017 to 18.03.2019. The program randomly selects a pair of cities from a predefined list of the largest cities in France and searches for available drivers. Trips start or end in Paris or its vicinity and have their other endpoints in one of the other 110 largest cities in France.

The program gathered all information accessible to prospective passengers. To do that, we open profiles of each driver available on a given route and collect all characteristics displayed on the profile, which include name, age, photo, a short biography, and the number of Facebook friends. Furthermore, we extract the entire history of received ratings and textual comments. We also observe the number of clicks and the number of sold seats for each listing. Clicking on the listing is necessary to book a trip, and a click opens a detailed description of the ride, but the passenger can still change her mind at no cost. We determine revenue by listing by calculating the product of the number of sold seats and

[13]The review system has a simultaneous reveal feature, which means that a user cannot observe a received review unless she has also posted one herself or the time to write one (two weeks) has elapsed. Only after both reviews have been sent do they become available to other users. Over the years, BlaBlaCar has introduced a few changes to the reputation system, which affected grading behavior. Appendix B discusses these changes.

price.

The listings that we observe have been featured on the platform for various periods of time. Some of them could have been posted just before our visit, while others could have been available for days. To account for this fact, we will control for how long a given listing is available and how many hours are left until departure.[14]

Additionally, we have matched our data with several other datasets. We establish gender and ethnicity using two complementary methods. First, we use the ethnic origins of names listed database published by the French government and supplemented with some other publicly available sources.[15] Second, we use a facial recognition software to improve our classification.[16] A detailed description of gender and ethnic identification is provided in Appendix C, where we show that both techniques - name and facial recognition - complement each other. Our definition of minority drivers is based on names with an Arabic or African origin or connotation; in doing so, we follow most of the existing literature. However, by considering both groups and using photo recognition together with name connotation, our approach improves the practice of assigning ethnicity compared to the prior studies in this context.

We proxy the quality of the car by approximating its value by the average price of the same type of car posted on eBay in Germany. The fuel efficiency of cars is calculated by matching car models with a dataset of long-distance fuel consumption of cars. We also collect data on city-level daily average fuel prices and highway tolls to construct instrumental variables for prices. Distances and expected travel time by car or public transportation are calculated for the moment of departure using Google Maps.

We also include information specific to destination and departure cities, such as population, median income, index of crime, and a share of foreign-born residents. Additionally, we have data on strikes related to transportation services (in particular, railways) that occurred in the spring of 2018. Descriptive statistics of selected variables are shown in Table 1. Appendix F lists the definitions of variables and sources of supplementary data.

One hundred eight thousand drivers appear in our dataset more than once. We use these observations to construct a panel. In the panel, the median number of observations per driver is five, and the mean is 12.

[14] This explains why many of our observations have zero sold seats and zero revenue. To check whether this biases our results, for a subset of our data, we have used the BlaBlaCar API to collect the final number of sold seats and revenue. We find similar results using this additional dataset.

[15] Translations of names with foreign origins into French exhibit considerable diversity. We phonetically encode our name lists and allow for minor spelling mistakes to improve our classification.

[16] www.kairos.com

Table 1: Descriptive statistics.

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Price (EUR)	552,518	31.43	15.98	6.00	18.00	41.50	78.50
Number of clicks	536,904	16.63	17.57	0.00	3.00	25.00	77.00
Sold seats	566,023	0.26	0.58	0.00	0.00	0.00	4.00
Revenue (EUR)	559,931	6.42	15.21	0.00	0.00	0.00	82.50
Minority	566,023	0.14	0.35	0.00	0.00	0.00	1.00
Male	552,530	0.73	0.44	0.00	0.00	1.00	1.00
Driver age	558,032	37.51	12.80	18.00	27.00	47.00	68.00
Number of reviews	560,331	37.12	60.71	0.00	4.00	42.00	421.00
Published rides (total)	537,681	38.84	49.29	0.00	7.00	50.00	256.00
Reputation	516,021	4.60	0.31	1.00	4.50	4.80	5.00
Seniority (months)	559,890	44.66	28.03	1.00	23.00	64.00	118.00
Posts per month	555,962	1.44	2.17	0.01	0.26	1.62	17.24
Photo	566,023	0.97	0.18	0.00	1.00	1.00	1.00
Bio (# words)	537,475	7.44	10.38	0.00	2.00	12.00	42.00
Car value (thousands of EUR)	471,117	6.08	5.04	0.60	3.10	8.06	24.40
Fuel consumption	486,604	5.00	0.77	3.65	4.39	5.39	7.50
Automatic acceptance	566,023	0.42	0.49	0.00	0.00	1.00	1.00
Hours until departure	508,754	95.50	107.47	0.001	20.96	126.47	501.69
Posted since	560,361	5.88	7.50	0.00	1.53	6.82	52.56
Travel time by public transport	545,200	3.97	2.42	0.14	2.25	5.41	15.24
Trip length (km)	550,118	396.34	192.27	67.32	232.00	491.68	906.46
Travel cost (fuel & tolls, EUR)	458,018	57.01	29.10	0.00	33.71	72.13	142.14
Train strike	566,023	0.04	0.19	0.00	0.00	0.00	1.00
Ride description (number of words)	509,243	13.49	14.60	2.00	2.00	22.00	93.00
Median revenue (city)	532,526	18.98	2.13	13.06	17.76	20.20	30.90
weekday	566,024	0.67	0.47	0	0	1	1
luggage size	116,982	0.89	0.31	0.00	1.00	1.00	1.00
detour	116,454	0.75	0.43	0.00	0.00	1.00	1.00
allows pets	223,774	0.22	0.41	0.00	0.00	0.00	1.00

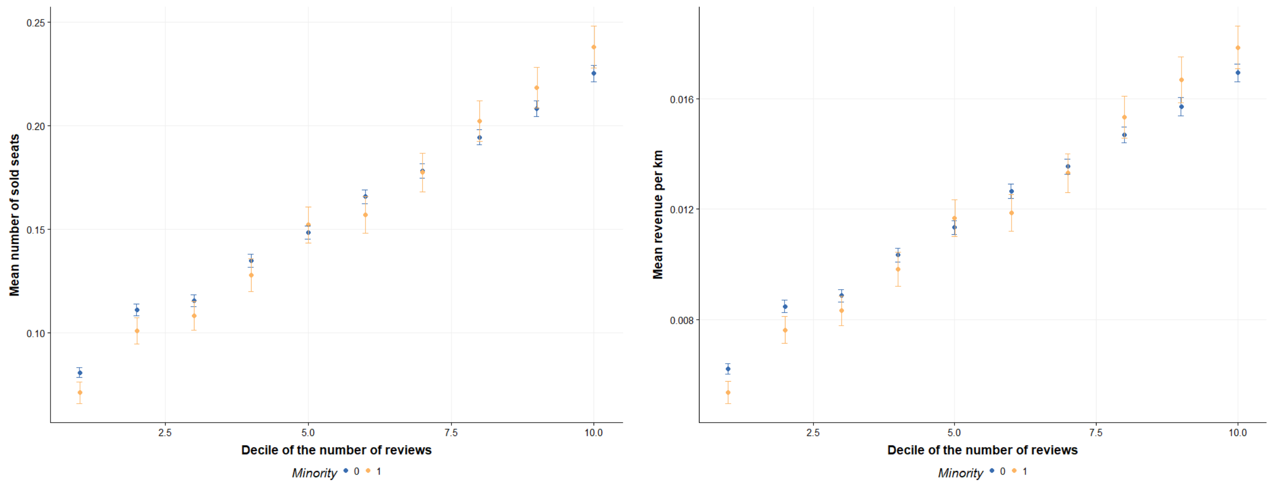
Note: See Appendix F for the definitions of variables and sources of supplementary data.

We have several measures of drivers' outcomes. First, the number of clicks is our proxy for the popularity of a listing. Passengers click on many drivers before deciding with which driver to travel. The mean number of drivers that a passenger can choose from is 30. The average number of clicks that a listing received is 17. The number of clicks is also useful for capturing the number of passengers searching for a ride in a given market.

Second, we observe the number of seats sold. On average, at the point of data collection, drivers managed to sell 0.3 seats. Drivers can change the price before the first passenger books a ride, but once one seat has been sold, the price remains the same. Hence, all passengers pay the same price. Third, the product of a price and the number of sold seats is revenue. In the structural model, we recover marginal costs; thus, we will also be able to measure economic profits.[17]

[17] Our dataset may miss some very successful rides that were no longer displayed when data were collected, which would lead to bias if the speed at which listings fill differs between minority and nonminority drivers. In Appendix E, we explore this issue and show that its magnitude is most likely not significant. However, as the most popular listings might be those of nonminority drivers, our estimates of the output gap should be regarded as a lower bound.

Figure 2: Gap between minority and nonminority drivers decreases with reviews



Note: Means of sold seats and revenue per kilometer within deciles of the number of reviews. Blue non-minority drivers; Yellow minority drivers.

3 Differences in outcomes across ethnicity and the impact of reviews

A quick look at the dataset reveals that minority drivers achieve lower outcomes than nonminority drivers. The raw data show that despite setting on average lower prices per passenger (31.9 EUR vs. 31.3 EUR), minority drivers receive fewer clicks (15.4 vs. 16.8), sell fewer seats (0.16 vs. 0.15), and in result earn lower revenue (3.45 EUR vs. 3.08 EUR).

In Figure 2, we compare outcomes across minority and nonminority drivers at different levels of reputation. We group drivers by ethnicity and deciles of the number of reviews and compute means of the number of sold seats and of the revenue per kilometer.

First, we observe that drivers that have more reviews sell substantially more seats and achieve higher revenue per kilometer. This effect has a strong economic significance with drivers in the tenth decile obtaining more than 2.5 times higher revenue than drivers in the first decile.

Second, we find that the difference between outcomes of minority and nonminority drivers changes with the level of reputation: amongst drivers that have just a few reviews, minority drivers are clearly underperforming as compared to their non-minority counterparts. The difference is statistically significant for drivers in the two first deciles. Across drivers with a higher number of reviews, ethnic differences are generally insignificant. Furthermore, across drivers with many reviews, minority drivers appear to outperform nonminority drivers; these differences are, however, statistically insignificant.

Averages presented in Figure 2, do not account for the fact that there are clear differences in routes in which minority and non-minority drivers operate, their cars, timing of trips etc., in Appendix G, we present the results of an OLS regression where we account for these differences. The general result carries over.[18]

Is the reputation effect due to selection? The evolution of the population of drivers on BlaBlaCar is characterized by frequent entries and exits. However, minority entrants are not more likely to quit than are nonminority entrants. The share of minority drivers is 14.6% among entrants, 13.2% in the intermediate group, and 15.6% in the experienced group. The share is relatively stable or even increasing, which suggests that selection cannot explain the reputation effect.

To provide further evidence that selection is not the mechanism behind the reduction of the disparity, in December 2018, we revisited profiles of drivers that appeared in our dataset earlier and collected their newly received reviews. The new data allow us to analyze usage intensity. We define two variables to measure the inactivity of drivers. Variable *exit* takes the value one if no new reviews were received between the last time a given driver appeared in the dataset and December 2018 and is zero otherwise. We also introduce a variable called *disaffection*, which takes the value one if the driver gathered fewer than five new reviews. Table 2 shows the results of the estimation of a logit model.

Table 2: Minority entrants are not more likely to exit the platform

	Dependent variable:	
	exit	disaffection
Minority	-0.129^{***} (0.028)	-0.097^{***} (0.030)
Entrant	1.350^{***} (0.024)	1.419^{***} (0.025)
Minority*Entrant	0.079 (0.065)	0.065 (0.066)
Age	-0.005^{***} (0.001)	-0.003^{***} (0.001)
Male	-0.098^{***} (0.018)	-0.084^{***} (0.019)
Seniority (number of months)	-0.005^{***} (0.0003)	-0.005^{***} (0.0004)
Posts per month	-0.731^{***} (0.010)	-0.736^{***} (0.011)
Bio (number of words)	-0.007^{***} (0.001)	-0.007^{***} (0.001)
Constant	-0.867^{***} (0.053)	-1.377^{***} (0.058)
Other driver characteristics	X	X
Time fixed effects	X	X
Observations	160,923	160,923

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

Note: Logit regressions, exit and disaffection as dependent variables.

First, minority drivers are more likely to continue using the platform. Second, new drivers are, generally, more likely to quit. However, we find no evidence that minority entrants are leaving the

[18] Additionally, we provide robustnesss checks in a panel setting, when we focus on drivers that have appeared in the dataset multiple times (see Appendix K) and a matching analysis (see Appendix J).

platform more frequently than nonminority entrants.[19]

These results suggest that the *reputation effect* is due not to a change in the composition of the sample, but to a causal impact of reviews. We do not observe, or model opportunity costs guiding drivers' entry and exit decisions. However, our findings are consistent with the idea that drivers are aware of the *reputation effect*. They realize that after a couple of periods of underachivement, their outcomes will improve; thus, they do not leave the platform despite facing initial discrimination: although the frequent exit of entrants is an essential aspect of the dynamics of the population of drivers on BlaBlaCar, the distinction between exit rates of minority and nonminority entrants is inconsequential.

3.1 Strategic behavior of drivers

Establishing reputation benefits all drivers, but is particularly valuable for minority drivers. In this section, we document how drivers respond to the incentive of acquiring a reputation. We explore two dimensions - efforts put into receiving higher grades and prices chosen by the drivers.

Arguably initial reviews are more consequential as they shift the posterior belief about quality to a larger extent. Therefore, if reviews reflect efforts exerted by drivers, the initial grades should be higher than the later ones - Figure 3 shows that this is the case. We restrict our attention to drivers who stayed on the platform at least until they obtained 30 reviews, and we explore the variation within their grades. Thus, survivorship bias does not influence the results. Figure 3 shows that drivers obtain, on average higher grades when they are new; the average grade decreases until the 10th review, at which point it stabilizes.

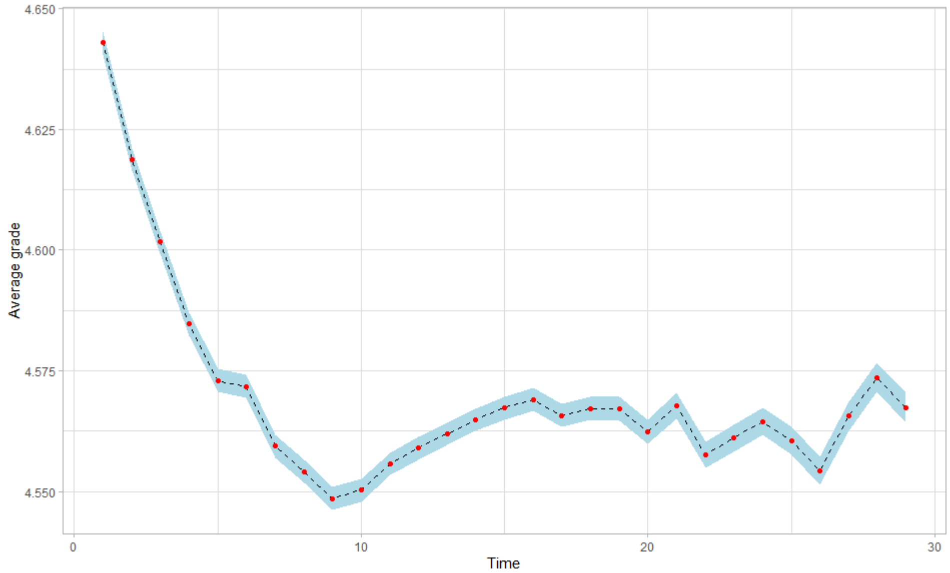
The extent to which the initial grades are higher varies across ethnic groups. In Figure 4, we show the difference between the early grades and the average grade a driver received after the 15th review. The gradual decrease in grades of minority drivers is more substantial than that for nonminority drivers. We interpret this as evidence of a larger effort that minority drivers exert to build a reputation.

Another way to boost reputation is to offer low introductory prices, which increases the chances of selling a seat and being reviewed. Minority drivers have an additional gain from accumulating reviews because they are, on average of higher quality than what the market expects. In Table 3 we present results of the estimation of within driver price variation.

We find that all drivers offer low introductory prices. The first few reviews lead to a significant increase in prices: the third review leads to an increase of 50 cents on average and the 5th review

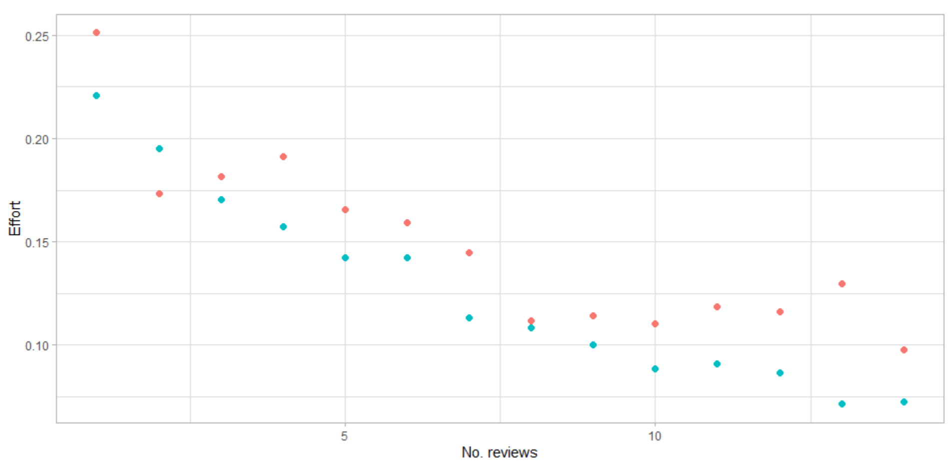
[19]The same analysis using the number of listings published (instead of the number of reviews collected) as a proxy for activity on the platform gives similar results.

Figure 3: Initial grades are higher



Note: The average grade from the first to 30th. Subset of drivers who used the platfrom at least until obtaining 30 reviews.

Figure 4: Minority drivers exert higher efforts



Note: Average early grades standardized by average late grades. Red dots: minority drivers, Blue dots: nonminority.

Table 3: Within driver price variation: the impact of reputation

	Dependent variable:	
	price	
	(1)	(2)
reviews:1-2	0.307 (0.221)	0.455^{*} (0.239)
reviews:3-4	0.495^{**} (0.235)	0.631^{**} (0.254)
reviews:5-8	0.691^{***} (0.241)	0.772^{***} (0.261)
reviews:9-12	0.910^{***} (0.260)	1.040^{***} (0.282)
reviews:13-16	0.798^{***} (0.280)	0.901^{***} (0.304)
reviews:17-20	0.857^{***} (0.303)	0.971^{***} (0.330)
reviews:1-2*minority		-1.040^{*} (0.632)
reviews:3-4*minority		-0.987 (0.676)
reviews:5-8*minority		-0.610 (0.688)
reviews:9-12*minority		-0.921 (0.731)
reviews:13-16*minority		-0.747 (0.783)
reviews:17-20*minority		-0.818 (0.835)
Ride controls	x	x
Driver FE	x	x
Day FE	x	x
Observations	78,903	78,903
R^2	0.658	0.658

$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Note: Within driver variation in prices, panel estimation. Reviews are binned and used as levels.

to an increase of 70 cents. However, there are decreasing returns from reviews. There is already no additional gain from the 9th review onwards. The last column of Table 3 introduces a distinction between minority and nonminority drivers. We observe that minority drivers set significantly lower prices when they have very few reviews; however, this effect disappears as soon as they have at least three reviews.

4 Railway strike as a quasi-experiment

Evidence presented in Figure 2 shows that minority drivers that are new to the platform receive higher boost to sales from each review. This allows them to gradually narrow the gap in outcomes. In this section, we exploit a natural experiment to argue that additional reviews cause higher sales and additional reviews for minority drivers, and particularly minority drivers with few reviews, have a stronger effect.

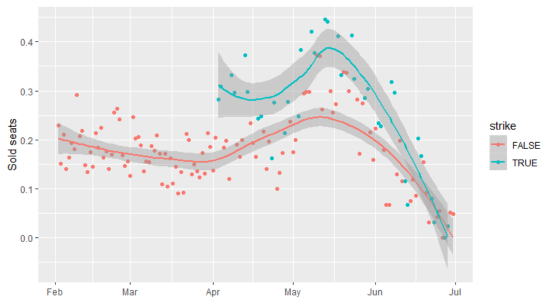
In the time-span of our data, French railway workers carried out a national strike.[20] The strike was organized as a sequence of two days of disruptions every five days for three months. BlaBlaCar and railways are in direct competition. A negative supply shock happening on railways transmits to BlaBlaCar as a positive demand shock. In April 2018, 5 million passengers traveled on BlaBlaCar, up

[20]Their opposition to plans to liberalize the European railway market and in particular to open the French market to competition was the main cause of the strike.

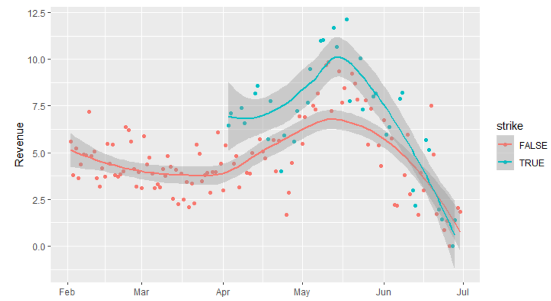
from an average of 1.5 million. The number of booking requests increased sixfold.[21]

All drivers, including minority drivers, faced significantly higher demand during the strike days. Figures 5A and 5B show an increase in the number of sold seats and revenue earned during the days of the strike.

Figure 5: Railway strike as a demand shock.



(A) Number of sold seats



(B) Revenue

Note: Horizontal axes time; red dots days without strike; blue dots days of strike.

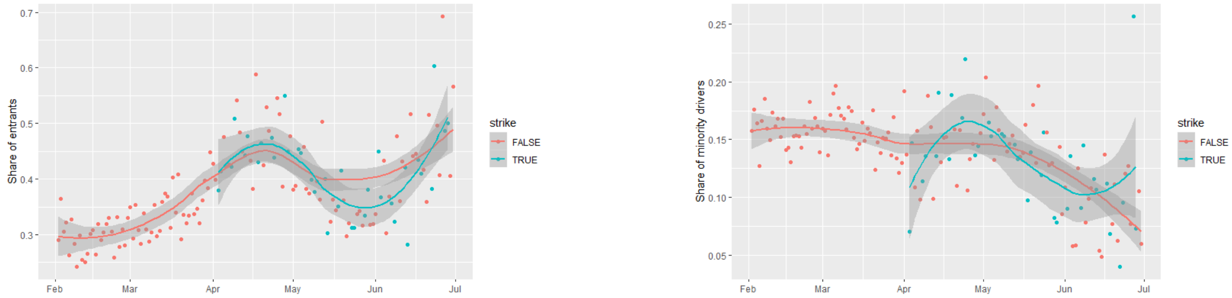
We interpret the strike as a natural experiment, where the treatment is a demand shock which translates to an exogenous increase in the number of reviews in the post-strike period. The critical assumption is that drivers did not select into treatment so that the increase in the number of reviews was exogenous. We argue that BlaBlaCar drivers are not professional drivers; they travel on a given route for other reasons and do not change their plans in response to a demand shock.

To support the assumption that treatment is exogenous to driver-specific characteristics, we compare the drivers on days with and without a strike. First, selection would result in an increased number of entrants traveling on the day of the strike. Figure 6A shows that there is no significant difference in the number of entrants on the days of strike and non-strike days. Second, minority drivers could be aware that it is easier to sell seats on a strike-day and be more inclined to post a ride. Figure 6B compares the share of minority drivers on strike and non-strike days; we do not observe increased entry of minorities on strike days. During strike days, 14.7% of drivers were minority drivers. On a non-strike day, in this period, the share was 14.8%. Table 20 in Appendix L compares other characteristics.

To show the impact of exogenous variation in the number of reviews on outcomes of ethnic minority drivers, we perform a difference-in-differences analysis. We follow the methodology developed

[21]Source: www.lemonde.fr/economie/article/2018/04/03/les-transports-alternatifs-grands-gagnants-de-la-greve-a-la-sncf_5279932_3234.html

Figure 6: No selection to treatment.



(A) Share of entrants **(B)** Share of minority drivers

Note: Horizontal axes time; red dots days without strike; blue dots days of strike.

by Sant'Anna and Zhao (2020). The estimator for the average treatment effect on the treated (ATT) is doubly robust, which means that it is consistent if either a propensity score or outcome regression models are correctly specified.

Treated drivers are those that happened to travel on a day of a strike. Pre-treatment period consists of rides that happened before the first day of strike, post-treatment period are trips that happened after the last day of the strike. We remove the period of strikes from the dataset. We treat our data as a repeated cross-section.[22]

We account for a rich set of covariates. Driver-specific information includes age, gender and reputation. Posting-specific covariates include the length of the trip, posted times, time until departure, time of day. Some other covariates are market-specific: number of available drivers, number of minority drivers, etc. Table 4 presents the results.

Table 4: Doubly robust difference-in-differences estimates with revenue as the dependent variable

Model	ATT	Std.Error	95% CI lower bound	95% CI upper bound
Minority with max 15 reviews	2.230	0.463	1.323	3.138
Non-minority with max 15 reviews	1.796	0.220	1.365	2.227
Minority drivers with 16+ reviews	0.195	0.669	-1.117	1.507
Non-minority drivers with 16+ reviews	-0.082	0.335	-0.738	0.574

Note: Estimates of the average impact of driving during a strike on revenue from trips after the strike period for four groups: minority drivers with 15 or fewer reviews during the strike period (row 1), non-minority drivers with 15 or fewer reviews during the strike period (row 2), minority driers with 16 or more reviews during the strike period (row 3), non-minority drivers with 16 or more reviews during the strike period (row 4).

[22] There are only a few drivers that we have observed in both pre and post-treatment periods that were treated.

Based on the evidence presented in Figure 2, we expect treatment effects to differ depending on ethnicity and the number of prior reviews. Thus, we consider four groups of drivers: minority drivers who are "entrants" (15 or fewer reviews), non-minority entrant drivers, "established" minority drivers (16 or more) reviews, "established" non-minority drivers.

We find that minority drivers that travelled during a strike receive an additionally revenue of EUR 2.23 in the trips after the strike, which corresponds to the difference between first and second decile from Figure 2. Non-minority "entrant" drivers also have a positive treatment effect, however, the magnitude is smaller EUR 1.796. We have more observations in the non-minority group, hence the tighter confidence interval. Both effects are statistically significant and the estimate of the effect for minority "entrants" does not fall into the confidence interval of the non-minority estimates, suggesting that the difference is statistically significant.

The estimates of treatment effects for "established" drivers both minority and non-minority are small and statistically not different from zero.

Driving during the strike allows drivers to sell out their seats and get reviews. Passengers on the subsequent trips appreciate this additional reviews and are more likely to purchase from these drivers than those that did not travel during the strike. The exogenous variation in the number of reviews suggest a causal relationship between a higher number of reviews and an additional revenue. Passengers reward minority drivers for additional reviews even more, the estimate of the average treatment effect is 24% higher.

5 Dynamic model of discrimination - demand and supply

We develop a model of discrimination in which sellers (drivers) compete in a market. Buyers (passengers) care about price and quality; pasengers learn about drivers' quality from group identity and past performance. Considering our empirical setting drivers belong to either a minority (m) or a majority (n).

5.1 Set-up of the model

Drivers. - Consider a driver who has observable group identity $g \in \{m, n\}$ and unobservable ability $\eta \sim N(\mu_g, 1/\tau_g)$, with mean $\mu_g \in \mathcal{R}$ and precision $\tau_g > 0$. Drivers have a cost of offering the service to a passenger equal to c. Drivers are active over a period of time $t = 1, 2, ...$ In every period, drivers set prices $p \in \mathcal{R}$ for their service and exert effort $a \in \mathcal{R}_+$; providing effort costs the driver $g(a)$, where $g(a)$ is increasing and convex. If a driver is trading in a given period she provides quality

$q_t = \eta + a_t + \epsilon_t$, where $\epsilon_t \sim N(0, 1/\tau_\epsilon)$ is an independent random shock with precision τ_ϵ. At time t Driver's problem writes

$$\max_{p_t, a_t} \left\{ \sum_{s=t}^{\infty} \delta^s (p_s - c) \mathbf{E}_t \left[\mathcal{S}_s(p_s, h_s) \right] - g(a_t) \right\} \tag{1}$$

,where $\mathcal{S}_s(p_s, h_s)$ is the number of sold seats in period s, which depends on the history of quality reports $h_s = (q_1, ..., q_{s-1})$, and $\delta \in (0,1)$ is a discount factor. We refer to $\pi_t \equiv (p_t - c)\mathbf{E}_t \left[\mathcal{S}(p_t, h_t) \right]$ as the expected driver's profit for a given price p_t and history of reviews h_t.

Passengers. - Passengers are active in one period; they observe available drivers, either pick one of them or decide not to trade. If they trade, they report quality q_t afterwards. Before choosing the driver, passengers observe drivers' groups g, histories of quality reports and prices p_t. Passengers' hold a prior belief that $\eta \sim N(\hat{\mu}_g, 1/\tau_g)$, $\hat{\mu}_g$ might not coincide with μ_g.[23]

Market. - In period t, there are M_t passengers, where M_t is a random variable such that $\mathbf{E}\left[M_t\right] = \mathcal{M}_t$, and N_t drivers. The market structure Ω_t summarizes the number of drivers and their characteristics (costs, histories of quality reports, and quality types). The market entry process is assumed to be exogenous and results in $\mathbf{E}\left[\Omega_t\right] = \mathbf{\Omega}$.

Timing. - The timing of the game is as follows: *i)* Drivers set prices that maximize their discounted sums of utility subject to histories of quality reports, costs, and expected market structure. *ii)* Passengers arrive to the market and observe available drivers. Each passenger either chooses the driver that maximizes her utility or decides not to trade, which gives a payoff of zero. *iii)* Drivers exert effort to maximize their discounted sums of utility. *iv)* Passengers observe quality and report it.

Equilibrium. - We assume that the set of potential drivers that can enter the market in a given period is large enough, so that drivers do not expect to compete against each other in subsequent periods. Thus, when setting p_t and a_t, drivers do not consider their impact on quality reports of other drivers.

[23] Alternatively each passenger j holds a belief such that $\mathbf{E}_t\left[\mu_g^j\right] = \hat{\mu}_g$.

5.2 Passengers' belief formation and choice problem

Belief formation and updating. - A passenger has a belief-based partiality toward nonminority drivers, if she believes that the average ability of nonminority drivers is higher than the average ability of minority drivers. Belief-based partiality can be biased or unbiased, depending on whether it coincides with the true population average for each group.

Definition 1. *A passenger has a belief-based partiality toward nonminority if $\hat{\mu}_n > \hat{\mu}_m$. This partiality is unbiased if $\hat{\mu}_n = \mu_n$ and $\hat{\mu}_m = \mu_m$, and otherwise is biased.*

Passengers learn about drivers' ability from the evaluation history. Their posterior belief is derived using Bayes' rule, given the prior belief on the average ability in group g. Passengers observe the quality q_t from the evaluation history. However, they do not distinguish the individual components (ability, effort, noise). To interpret the evaluation history passengers need to form their belief about the level of effort exerted by the driver in the past. Let a_t^* be an equilibrium level of effort in period t. Thus, from the quality report q_t passengers' learn $z_t \equiv q_t - a_t^* = \eta - \epsilon_t$ and the posterior belief about the driver's ability writes

$$\mathbf{E}_{t+1}\left[\eta | h_t\right] = \frac{\tau_g \hat{\mu}_g}{\tau_g + t\tau_\epsilon} + \frac{\tau_\epsilon}{\tau_g + t\tau_\epsilon} \sum_{s=1}^{t} z_s. \tag{2}$$

Increasing the precision of the quality reports (or decreasing the randomness of the outcome of the task) enlarges the weight assigned to the past performance, while increasing the precision of the distribution of ability boosts the importance of the prior belief.

Passengers' choices. - A passenger j chooses between drivers and the outside option of not trading to maximize her utility u_{ijt} ; the utility depends on quality (positively on the expectation and negatively on the variance), price, and passenger-driver- specific shock ε_{ijt}, which is independent across drivers and distributed following type-I GEV distribution,

$$u_{ijt} = \alpha \mathbf{E}_t\left[q_{it} | h_{i,t-1}\right] + \beta \mathbf{Var}_t\left[q_{it} | h_{i,t-1}\right] + \gamma p_{it} + \varepsilon_{ijt} \tag{3}$$

, where α, β, and γ are the marginal values of expected quality, the variance of quality, and income respectively. Thus, the passenger j chooses driver i, when

$$\mathbf{E}_t\left[u_{ijt}|h_{i,t-1}\right] = \max\left\{\mathbf{E}\left[u_{kjt}|h_{k,t-1}\right],0\right\}\forall k\in N_t. \tag{4}$$

Discrimination. - Discrimination is the disparate treatment of drivers based on the group to which the driver belongs, rather than individual attributes. In our framework, a passenger discriminates against minority drivers, when a minority driver is chosen with a lower probability than a nonminority driver with the same price and history of quality reports. Let

$$D(h,p) \equiv \mathbf{E}\left[\mathcal{S}(p,h)|n\right] - \mathbf{E}\left[\mathcal{S}(p,h)|m\right] \tag{5}$$

denote the difference between the expected number of sold seats of a nonminority and minority driver conditional on the same histories of quality reports and prices. Note that $D(h,p)$ does depend on the number of reviews t trough the length of history h.

Definition 2. *A minority (nonminority) driver faces discrimination if $D(h,p) > 0$ ($D(h,p) < 0$).*

In our framework, discrimination is the property of behavior of passengers, while belief-based partiality is a property of the primitives of the model.

5.3 Dynamics of efforts

Drivers' effort is noncontractible and it is exerted after passengers choose drivers; the incentive to exert a nonzero level of it is driven by the impact of quality evaluation on future profits. The first order condition of drivers' maximization problem writes

$$\sum_{s=t}^{\infty}\delta^{s-t}\mathbf{E}\left[\frac{\partial\pi_s}{\partial a_t}\right] - g'(a_t) = 0. \tag{6}$$

To obtain the utility maximizing level of effort a driver equates the marginal benefit, which is the increase in future profits, with marginal cost, the derivate of the cost of effort function.

Proposition 1. *The equilibrium sequence of effort $\{a^*\}_t$ tends towards zero as the driver gains experience: $\lim_{t\to\infty} a_t = 0$.*

The proof of Proposition 1 is provided in Appendix M.1. In the proof, we first use the consumer choice problem to rewrite drivers' profits. Second, we consider the limit of optimal effort as t goes to infinity. Proposition 1 generalizes the main result of Holmstrom (1999) by allowing elastic demand and considering the role of strategic pricing by drivers.

Drivers' exert effort to increase future profits. Since initial reviews have a substantial impact on the posterior beliefs, the level of effort is high when the number of reviews is low. As more reviews become available, the residual uncertainty about the driver's type tends to zero, thus the incentive to exert effort tends to zero as well.

The level of effort depends on the impact of effort on the future profits which writes (see Appendix M.1 for details)

$$\frac{\partial \mathbf{E}\left[\pi_s\right]}{\partial a_t} = M_s \mathcal{S}_{is}(1-\mathcal{S}_{is})\alpha \frac{\tau_\epsilon}{\tau_{gs}}(p_{is}-c_i). \tag{7}$$

The change in the ratio of informativeness of a review to the remaining uncertainty (τ_ϵ/τ_{gs}, where $\tau_{gs} = \tau_g + s\tau_\epsilon$) generates the negative trend: with each review τ_{gs} increases, thus the driver exerts less effort. When the quality is observed with substantial error (or the quality of the service itself is realised with high randomness) the level of effort is lower. Also, large variance in the distribution of ability τ_g, results in high uncertainty of the quality realisation and thus forces the driver to exert higher effort. A higher expected quality is more valued when the elasticity of demand with respect to quality (α) is large.

Driver's expectation of the future number of sold seats is the other key determinant of the level of effort. M_s shifts the level of effort up, while the term $\mathcal{S}_{is}(1-\mathcal{S}_{is})$ is increasing up to the moment when the probability of selling a seat to each passenger is one half. If the driver expects to sell a high (low) number of seats, she exerts a larger (smaller) effort.

Corollary 1. *The equilibrium level of effort of minority drivers:*

1. *Decreases, relative to nonminority drivers, when the discrimination against minority drivers $D(h,p)$ increases,*

2. *Is higher, when discrimination is associated to a negatively biased belief-based partiallity toward nonminority ($\hat{\mu}_m < \mu_m$) than an unbiased belief-based partiality ($\hat{\mu}_m = \mu_m$).*

The proof of Corollary 1 is in Appendix M.2. This result shows how passengers' beliefs and resulting discrimination influence drivers' incentives to exert effort. Exerting effort entails costs that are sunk in the current period, hoping to recoup them by increased future profits. When a driver is expecting to sell many seats in the future, thus can raise the price on a higher market share, the returns from such an investment are higher.

When passengers' believe that the mean ability amongst minority drivers is low, they will buy fewer seats, reducing minority drivers' incentive to exert effort. However, if the belief is negatively biased so over time the number of sold seats will increase, the driver will take that into account exerting higher effort than in the case of unbiased belief-based partiality.

5.4 Dynamics of pricing

While exerting effort increases grades, changing the price influences the probability of selling, and thus of receving a grade at all. Increasing the number of grades influences the probability of selling through two channels: reducing the variance and influencing the expected grades.

Proposition 2. *The equilibrium sequence of prices* $\{p^*\}_t$ *tends towards* $\tilde{p}$*, where*

$$\tilde{p} \equiv \arg\max \{(p-c)\mathbf{E}[\mathcal{S}(p,\eta)]\} \tag{8}$$

is the profit maximizing price under complete information, where $\mathcal{S}(p,\eta)$ *is the market share of a driver who has an infinite history of quality reports* $h_t = \eta + a_t^*$ *for all* t*.*

Proof of Proposition 2 is in the Appendix M.3. In the proof we show that the limit of expected quality, given the equilibrium level of efforts, is driver's true type η and that the conditional variance shrinks to variance of the quality reports τ_ϵ.

Corollary 2. *As a driver receives quality reports her price sequence:*

- *Increases towards* $\tilde{p}$*, when* $\mu_g < \eta_i$*,*
- *Either increase or decrease towards* $\tilde{p}$ *otherwise.*

Two effects control the evolution of prices, decreasing variance increases the prices, while the quality updates can either increase or decrease depending whether μ_g is higher or lower than the η_i. In the empirical section later, we show that the dynamic pricing exacerbates these effects compared to static prices, this is because if a driver anticipates to have her quality revised upwards she will decrease initial price, and if she expects a downward revision she increases the current price.

Corollary 3. *The equilibrium level of introductory prices of minority drivers decreases when the difference between the belief about the mean ability and the mean ability* $|\hat{\mu}_g - \mu_g|$ *increases.*

This observation indicates that the minority drivers have an incentive to *invest* in reputation, by offering low introductory prices. When the driver's expectation of the grade is higher than that of the market, the driver has an incentive to reduce the price in order to increase the probability of selling and benefiting from having the market beliefs revised upwards in the next period. The larger the biased belief-based partiality the higher the incentive to reduce the introductory price.

5.5 Dynamics of discrimination

In our framework, discrimination arises due to a combination of incomplete information and the belief that mean ability differs across groups ($\hat{\mu}_m \neq \hat{\mu}_n$). The impact of the beliefs about the group mean ability is gradually loosing importance as drivers gain reviews. Proposition 3 formalizes it.

Proposition 3. *As drivers gain quality reports discrimination tends to zero*

$$\lim_{t \to \infty} D(h, p) = 0. \tag{9}$$

Proof of Proposition 3 is in Appendix M.4; it is a direct consequence of the beliefs updating via Bayes Rule and the efforts following equilibrium sequence $\{a^*\}_t$. Note, that the outcomes of individual drivers might diverge as drivers receive reviews; however, conditioned on these reviews the outcomes converge. Discrimination, in our framework, is due to incomplete information, as drivers are evaluated based on their individual qualities rather than on group means.

Quality reports revise the beliefs held by the market. Thus, when these beliefs are biased there is a change in the number of sold seats by the drivers on average. Let $\mathcal{A}_{g,t} = \sum_{i=1}^{N^g} \mathcal{S}(p, h)$ be the average number of sold seats by drivers in group g with t quality reports ($t = Card(h_t)$), where N^g is the number of drivers from group g, and $\Delta(\mathcal{A}_t) = \mathcal{A}_{n,t} - \mathcal{A}_{m,t}$ be the difference of these averages.

Corollary 4. *Dynamics of the average number of solds seats* $\mathcal{A}_{g,t}$ *depend on the form of belief-based partiality:*

- *Under unbiased belief-based partiality* $\Delta(\mathcal{A}_t) = \tilde{\Delta}$*, which does not depend on the number of quality reports, beyond their impact on the decrease in variance,*
- *Under biased belief-based partiality* $\lim_{t \to \infty} \Delta(\mathcal{A}_t) = \tilde{\Delta}$*, and when the bias is negative* $\Delta(\mathcal{A}_t) > \tilde{\Delta}\ \forall t$*.*

The bias in the beliefs is gradually corrected by the quality reports and it tends to the diffference in outcomes that can be justified by the difference in underlying ability. Thus, if the difference in outcomes tends to the differenece in types. When the belief is overly pessimistic it gradually increases, if the overly negative it gradually decreases.

6 Identification and estimation

In this section, we present assumptions under which we can identify parameters of interest and estimate them. There are generally three groups of parameters. First, demand elasticities α, γ, β and θ. The key observables are prices and the numbers of sold seats. Additional available information is provided by the conditioning variables r and X for all drivers in each market. The second group consists of parameters related to the model of belief formation and updating, where we are interested in the prior beliefs μ_m, true distributions of types $\mathcal{N}(\hat{\mu}_m, 1/h_m)$ in each population m, and the informativeness of the reputation system h_ϵ. The observables that we will use to recover these parameters are the histories of grades of individual drivers and their market outcomes. Finally, the third group contains supply-side parameters including drivers' types η_i, efforts a_{imt}, and marginal costs c_i. We will also identify the cost of effort function $g(a_{imt})$. The observables of the supply-side involve the prices set by drivers, the histories of their grades, and transition probabilities from one history set to another, defined formally later.

6.1 Demand estimation

We propose a standard conditional logit model of demand. Here, we discuss some of its main features; detailed proofs and further discussion are provided in McFadden (1974). We assume that the utility of passengers is linear in the characteristics of drivers, that is,

$$u_{ijtm} = \alpha \mathbf{E}\left[w_{itm}|w^{it}\right] + \gamma p_{it} + \beta r_{it} + \mathbf{X_{it}}\theta + \varepsilon_{ijt},$$

where subscript j refers to passengers. In our baseline model, the stochastic term ε_{ijt} is the only difference between the passengers. We assume that it is a random variable with an extreme distribution $\mathcal{F}(\varepsilon_{ijt}) = \exp(-\exp(-\varepsilon_{ijt}))$. The probability that passenger j chooses driver i from N available drivers (indexed by k) and the outside option is $P_{ij} = P(u_{ij} \geq u_{ik}, \forall_{k\neq i})$, which given the assumption on u_{ijtm}, is

$$P_{ij} = \frac{\exp\left(\alpha \mathbf{E}\left[w_{itm}|w^{it}\right] + \gamma p_{it} + \beta r_{it} + \mathbf{X_{it}}\theta\right)}{1 + \sum_{k=1}^{N} \exp\left(\alpha \mathbf{E}\left[w_{ktm}|w^{kt}\right] + \gamma p_{kt} + \beta r_{kt} + \mathbf{X_{kt}}\theta\right)},$$

where the utility of the outside option is normalized to zero. McFadden (1974) shows that the log-likelihood function with these choice probabilities is globally concave in the parameters of demand.

Thus, we can estimate its parameters by maximizing likelihood function with M observations (passengers),

$$\max_{\alpha,\gamma,\beta,\theta} \sum_{j=1}^{M} \sum_{i=1}^{N} d_{ijt} \ln P_{ij}(\alpha, \gamma, \beta, \theta)$$

,where $d_{itj} = 1$ if passenger j chooses driver i, and $d_{itj} = 0$ otherwise.

The identifying assumption is that our controlling variables X_{it} capture all demand-relevant driver-specific characteristics so that there is no heterogeneity across drivers that is observed by passengers but not by us. We make this assumption because, in our dataset, we indeed observe all information that is available to passengers. Nevertheless, for robustness, we introduce instrumental variables (cost shifters) in Appendix O to control for potential endogeneity, and also introduce random coefficients.

The number of potential passengers M is measured directly in our dataset. We have previously used the number of clicks that each listing received to measure the respective listing's popularity. However, the total number of clicks in the market can proxy the number of potential passengers. Within a market, defined as a route-and-day combination, we use the highest number of clicks received by any listing to represent the total number of passengers that have been interested in booking a ride. The difference between the maximum number of clicks and the total number of sold seats proxies the number of passengers that have searched for a ride, but did not buy. In other words, the latter passengers chose their outside option. The market size measured in this way exhibits significant time variation.

Market prior beliefs: We do not observe the market's belief about the expected quality $\mathbf{E}\left[w_{itm}|w^{it}\right]$. However, we know that passengers' beliefs converge to underlying quality as drivers receive reviews. Thus, drivers who have accumulated a substantial number of reviews face correct beliefs, which are consistent with their observed reputation. To recover market beliefs about drivers with no or few reviews, we will first estimate demand using a subset of markets where there are only experienced drivers (10 thousand out of 60 thousand markets).

In the second step, we use the estimated demand to predict the expected number of sold seats for the entire dataset. If, for a subset of drivers (for example, minority drivers), passengers at the booking stage are systematically incorrect about the grade they will give after the ride, the predicted market share obtained with our model will differ from the observed number of sold seats. We will use this difference to obtain the disparity between the grade given after the trip and the market expectation of a grade. To do that, we compare the market outcome s_{imt} to the prediction and assign the entire prediction error to passengers' errors in the assessment of the expected quality $\tilde{w}_{imt}$:

$$\tilde{s}_{imt} - s_{imt} \propto \tilde{w}_{ijt} - \mathbf{E}\left[w_{imt} | w^{imt}, \mu_m\right], \tag{10}$$

where $\mathbf{E}\left[w_{imt}|w^{imt}, \mu_m\right]$ is the market's belief about the expected quality of driver i, from population m with a history of grades w^{imt}.

Furthermore, from the model of belief formation and updating, we obtain a functional form of the expected quality. We attribute the difference to the disparity between the belief about the mean type in population μ_m and the actual mean $\hat{\mu}_m$.

6.2 Supply-side parameters

The key supply-side parameters are drivers' types η_i, their efforts a_{imt}, and marginal costs c_i. For all drivers in our dataset, we have histories of ratings obtained from the driver's first ride until the moment the driver appears in our dataset for the last time. We will use these grades to recover drivers' types and efforts.

Figure 3 of section 3.1 shows the average ratings at different stages of drivers' careers.[24] We observe that the ratings are high in the beginning and stabilize as more reviews become available. The observed trajectory of grades is consistent with the prediction of the model - the initial increase in grades is due to efforts, while the level at which the grades stabilize coincides with the driver's type.

By Proposition 1, the optimal level of effort approaches zero as t tends to infinity. We assume a burnout period t^*, after which the level of effort is low.[25] Thus, we define the parameters of interest as follows:

- The intrinsic quality (type) of an individual driver is the average of her grades after t^*,

$$\eta_i = \frac{\sum_{t=t^*}^{T} w_{imt}}{T - t^*}, \tag{11}$$

where T is the last period in which we observed a grade given to driver i.

[24] The first point on the left chart is the average first grade. We restrict the sample to drivers who stayed on the platform long enough so that they gathered enough reviews to reveal their types. Restricting the sample has an additional advantage of mitigating the survivorship bias stemming from the selection of the drivers with high grades. As pointed out in section **??**, receiving a low grade increases the chance of a driver leaving the platform; thus, the grades of drivers who stayed were on average higher than the ratings of those who left the platform early on.

[25] In practice when it is no longer statistically significant for both minority and nonminority drivers.

- The effort a^*_{imt} of driver i from population m at time t with history of grades w^{it} is

$$a^*_{imt} = \frac{\sum_{s=1}^{N^{m,w^{st}}} (w_{smt} - \eta_s)}{N^{m,w^{st}}}, \tag{12}$$

  where s indexes drivers from population m with history of grades w^{st}, and $N^{m,w^{st}}$ is the number of such drivers in our dataset. Thus, the expected effort of driver i is the average difference between grades and types for all drivers with the same characteristics (including the number of reviews), types, and histories of grades.

- We assume that the distribution of the error term is normal with zero mean. We are interested in estimating the precision (the inverse of variance) of the error term, which is given by the inverse of the mean of variances of grades after t^*,

$$h_\epsilon = \frac{N^{t^*}}{\sum_{s=1}^{N} Var(w_{st})} \forall_{t>t^*}, \tag{13}$$

  where N^{t^*} is the set of grades of drivers with $t > t^*$.

We need several assumptions to identify these parameters in the data. First, there are no listing-specific variables other than types, efforts and exogenous errors that influence grades. In particular, we assume that prices do not influence grades. Appendix N provides some evidence supporting this assumption. Second, error terms are random variables, with mean zero. We require that: $\mathbf{E}[\epsilon_{it} + \epsilon_{it+1}] = \mathbf{E}[\epsilon_{it}] + \mathbf{E}[\epsilon_{it+1}]$. This is necessary, so that

$$\lim_{T\to\infty} \left[\frac{1}{T - t^*} \sum_{t=t^*}^{T} (\eta_i + \epsilon_{it}) \right] = \eta_i.$$

Having determined the types of individual drivers, we obtain the distributions of types in different populations. The mean is given by

$$\hat{\mu}_m = \frac{1}{N^m} \sum_{s=1}^{N^m} \eta_s,$$

and the precision is

$$h_m = \frac{N^m}{\sum_{s=1}^{N^m} (w_{smt} - \eta_s)^2},$$

where N^m is the number of drivers in population m.

Estimation of the cost of effort function: The cost of effort function is unknown to us, we have assumed that it is convex and increasing. The function $g(\cdot)$ defines the optimal level of effort by equating the marginal benefit from exerting a unit of effort with the cost of such a unit.

$$a_{imt} = g^{-1'}\left(X_i^1, X_{it}^2\right). \tag{14}$$

, where X_i^1 are driver's time invariant characteristics and X_{it}^2 are their time variant features. Together X_i^1 and X_{it}^2 characterize driver's current and future profits. From the discussion above, we know how to measure the levels of effort. Thus, we can approximate function $g(\cdot)$, to do that we split the dataset into train and test and try various functional forms. Table 5 presents the results.

Table 5: Performance of models predicting effort

model	dataset	R2	MSE	std.error
Constant	train	<0.001	0.363	0.001
Constant	test	<0.001	0.360	0.002
LASSO	train	0.002	0.362	0.001
LASSO	test	0.002	0.359	0.002
GBM	train	0.023	0.354	0.001
GBM	test	0.010	0.356	0.002

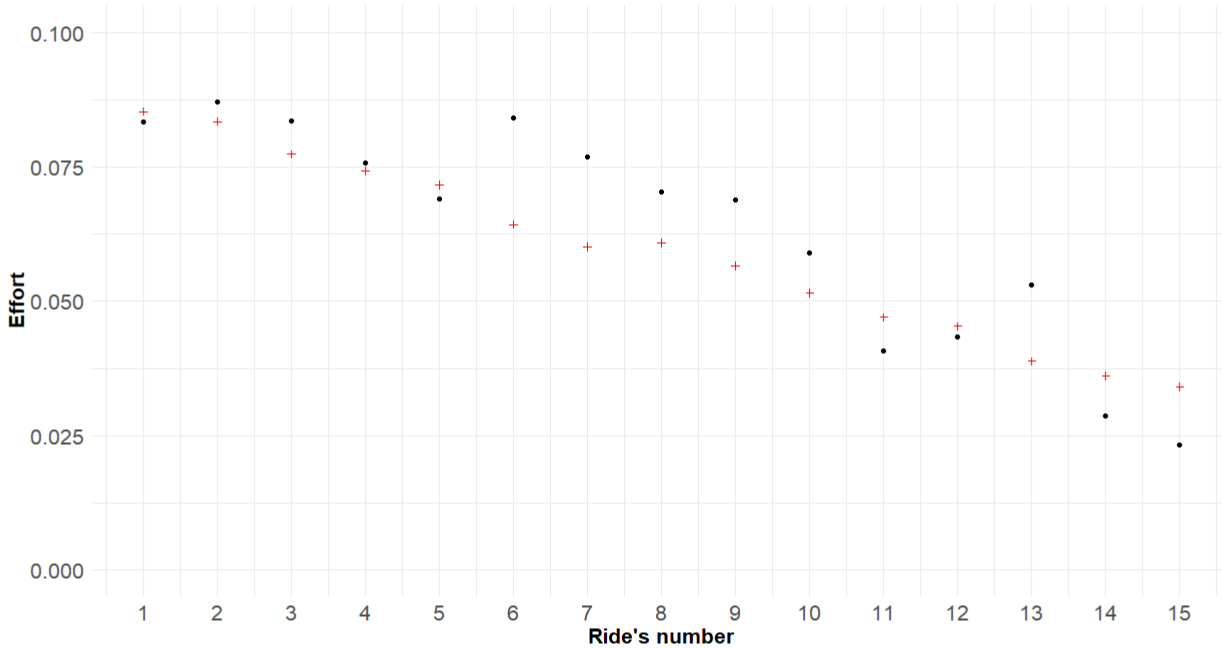
Note: Constant model is a model with just intercept. Covariates in LASSO are the number of reviews, reputation at the time of the trip, driver's age, months since joining, picture, length of the bio, gender, price of the car, automatic acceptance, and minority. Covariates are fully interacted. GBM takes the same set of covariates. All models trained on 70% of data and tested on the remaining 30%.

The first two rows of table 5 present the performance of a model with only a constant, rows three and four show the LASSO's model performance; LASSO includes a fully interacted set of drivers' covariates. The last two rows report results from GBM (Gradient Boosted Machine) with the same set of covariates as LASSO.

We can notice from table 5, that drivers' covariates explain only a small portion of the variation in efforts, part of this depends on passengers' tastes which we do not observe. Nevertheless, we can observe that both LASSO and GBM use driver's covariates to explain a part of the variation. GBM has the best performance in the test set and we will proceed with this model.

Figure 7 shows the observed and predicted efforts (prediction of GBM model on a test set).

Figure 7: Observed and predicted effort across driver's number of reviews



Note: Horizontal scale the rank of the trip. Vertical axis effort - blue dots observed means, red crosses: means of predictions. Prediction based on the gbm model trained on 70% of data, prediction on the test set.

We can notice that the predicted effort decreases with the number of reviews that the driver already has.

Pricing stage

The marginal cost of having a passenger on board defines the profitability of using the platform. We argue that drivers act strategically while setting their introductory prices, and in section 3.1, we provide some evidence of this. In the reduced-form results, we also show that the returns from reputation are decreasing. There is a saturation point at approximately ten reviews, after which there are no more incentives to reduce prices to receive more reviews. Thus, prices in periods from the first to the tenth exhibit the static-dynamic tradeoff, while prices in periods after the tenth can be interpreted as static profit-maximizing prices.

Assuming that after the tenth period, prices maximize static profits, we can recover marginal costs. To this end, we need to first estimate markups, which we obtain from estimated demand. The difference between the price and the markup for prices of drivers with more than ten reviews is given by

$$p^*_{imt} = c_i + \frac{s_{imt}}{\frac{\partial s_{it}}{\partial p_{imt}}}. \tag{15}$$

In this way, we recover marginal costs and obtain their distribution for drivers who stayed at least until the tenth period.

As we argued while discussing the model, each price is a solution of a dynamic programming problem. We observe the transition rules $p(w^{imt+1}|w^{imt})$ and $p(X'_{imt+1}|X_{imt})$ directly in the data; $p(w^{imt+1}|w^{imt})$ is the probability of obtaining each possible grade conditioned on having the level of reputation w^{imt}, and $p(X'_{imt+1}|X_{imt})$ is the probability of receiving a grade after selling a seat.

To find the optimal price in period t, we need to first characterize the optimal behavior in period $t+1$, because the value of being in period $t+1$ defines the incentive to get there. Hence, to solve the problem we will proceed by backward induction. First, in period ten, we assume Bertrand pricing. We find the optimal price for a driver with a given set of characteristics and a marginal cost. We also obtain the value of being in period ten (the discounted sum of profits). Then, in period nine, there is already an incentive to reduce the price to proceed faster to period ten. Hence, the problem in period nine is written as

$$\begin{aligned} p^*_{im9} = \arg\max\{ & (p_{im9} - c_i) M_9 s_{im9}(p_{im9}, X_{im9}) \\ & + \delta \left[s_{im9} V_{im}(10) + (1 - s_{im9}) \left((p_{im9} - c_i) M_9 s_{im9} + \delta (s_{im9} V_{im}(10) + (1 - s_{im9})(p_{im9} - c_i) M_9 s_{im9} + ...) \right) \right] \}, \end{aligned} \tag{16}$$

where $s_{im9}(p_{im9}, X_{im9})$ is the probability of selling a seat given the price p_{im9} and characteristics X_{im9}, $M_9 s_{im9}(p_{im9}, X_{im9})$ is the expected number of sold seats, which determines the expected number of new reviews, and $V_{im}(10)$ is the expected value of being in period ten (expected with respect to the grade that i will obtain). If driver j does not sell a seat, she solves the problem of period nine again until she obtains a review.

After determining the optimal price for period nine, we proceed to period eight and so forth until we reach the task of determining introductory prices. Any price that we observe is the solution to this problem; thus, we can identify the marginal cost from each observed price.

7 Results

In this section, we present and discuss the results of the estimation of the model. First, we show estimates of demand. Second, we demonstrate how incentives to exert effort differ across minority and nonminority drivers. Next, we present the estimated prior. Finally, we show estimates of marginal costs and discuss how the incentives to invest in reputation depend on them.

Demand estimates

In Table 6, we present the results of demand estimation. The dependent variable is d_{ijt}, a binary variable that takes the value one if driver i was selected by passenger j and is zero otherwise. As we proceed from column one to column two, we add more controls. The variable *type* is the average grade from the tenth onwards, while *reputation* (column three) takes into account all grades available on drivers' profiles. The regression presented in column one controls for the type, the number of reviews, and price; in column two, we add a full set available controls, time, and trip specific effects. In column three, we use *reputation* instead of *type*. Demand is estimated using a subset of 10241 markets (400 thousand choice situations). We will use model two in the supply-side estimation and the analysis of counterfactuals.

Table 6: Demand estimates

	Model 1	Model 2	Model 3
Ride price	-0.00 (0.00)***	-0.00 (0.00)***	-0.00 (0.00)***
Type	0.12 (0.06)**	0.13 (0.06)**	
Log(number reviews)	0.15 (0.02)***	0.14 (0.02)***	0.15 (0.02)***
Minority		0.06 (0.05)	
Reputation			0.24 (0.10)**
AIC	31929.66	30150.03	30145.51
R^2	0.45	0.45	0.45
Max. R^2	0.49	0.49	0.49
Num. events	154259	147905	147905
Num. obs.	470165	442839	442839

Note: Demand estimates: subset of markets. All coefficients presented in the Appendix P For additional robustness check we also estimate a model with random coefficient associated with price (BLP); elasticities of price and quality are in the Appendix O.

Demand is generally not very elastic. The marginal value of income and expected quality is -0.12 and 0.57, respectively

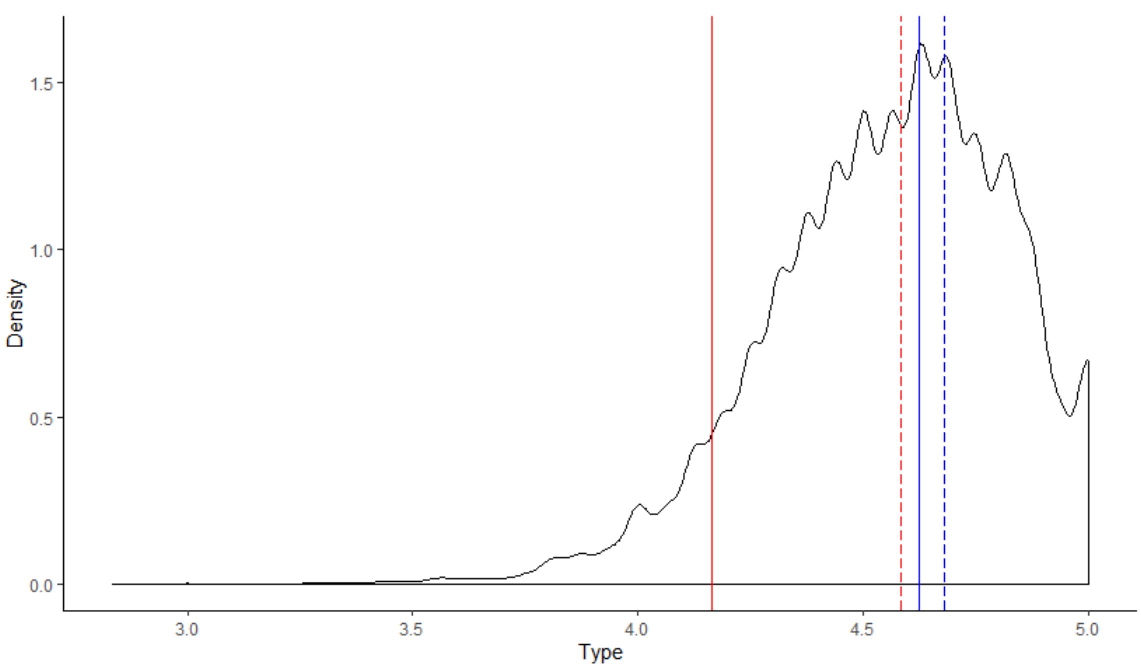
Market prior beliefs

To recover passengers' beliefs about the expected quality of service offered by drivers with no reputation, we first predict the number of sold seats using the *model 2* fromTable 6. Next, we attribute the error of the prediction to the expectation of the grade.

We find that minority drivers with no reputation are expected to deliver the quality of 4.16 (on a scale of 1 to 5), while they are graded at 4.619 on average. The expectation corresponds to the 7.5th percentile of the distribution of quality. For comparison, nonminorities are expected to provide a quality of 4.59 and are graded 4.68. Figure 8 summarizes this. The solid blue line represents the average first grade obtained by minority drivers, while the solid red line is the market expectation of the grade. Dotted lines correspond to grades and their expectations for nonminority drivers. The distribution of all grades is shown in black.

Finally, as argued throughout this paper, the beliefs about quality are being updated; thus, the two numbers converge. Minority non-entrants (with more than two reviews) are believed to be of quality 4.539 before the trip and are graded 4.592 ex post.

Figure 8: Erroneous beliefs and given grades



Note: The distribution of grades is shown in black. Blue lines represent the mean first grade obtained by a minority driver (solid line) and by nonminority drivers (dotted line). Red line illustrate the market beliefs on the expected quality (minority- solid line, nonminority - dotted line).

The expected quality in the first period depends only on the prior belief about the distribution of quality among minority drivers and the expected level of effort. Consistently with the result of Corollary **??**, minority drivers exert greater effort than the market expects them to. Given the estimated parameters, the difference between the two levels of effort results in the difference of 0.04 in the first grade.

The distribution of types in the population of nonminorities has mean of 4.56 and variance 0.07. For the population of minority drivers, the mean is 4.49, and the variance is 0.09. We account for the difference in expected and exerted efforts and find that the market expects the mean type of minority drivers to be at 4.09.

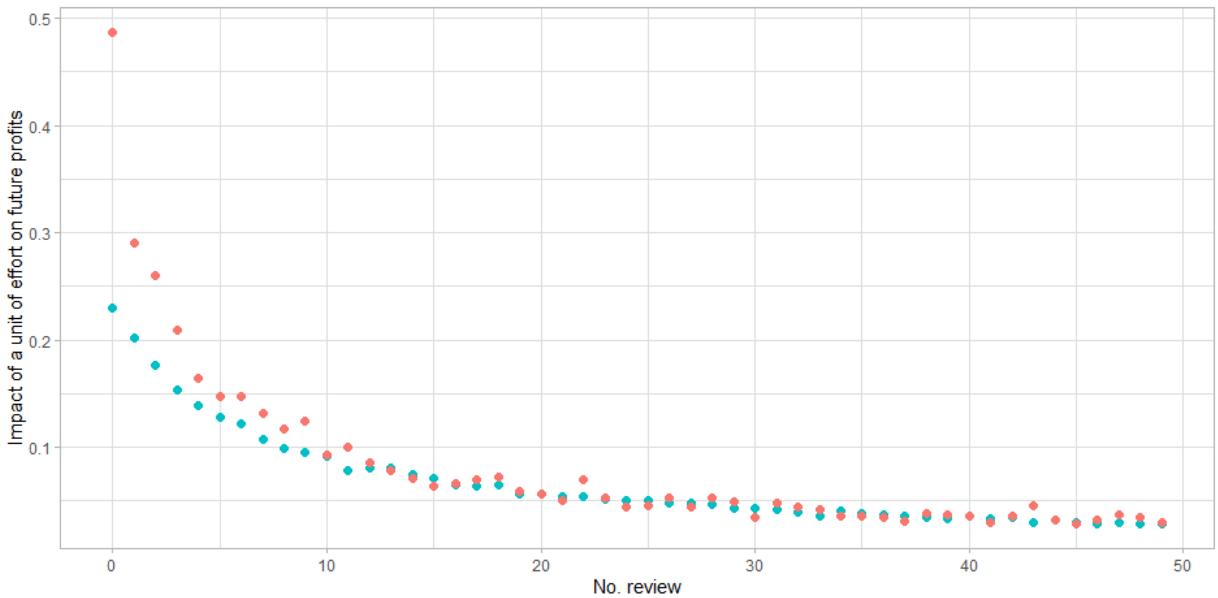
Incentives to exert effort

The incentive to exert effort is determined by the magnitude of the impact of a unit of effort on future profits. Figure 9 shows the average (across all drivers) increases in the next period's profits due to a unit of effort exerted in the current period, determined as $\frac{h_\epsilon}{h_{mk}} \frac{\alpha}{\gamma} \mathbb{E}[M_k s_{ik}]$. We show how this quantity changes during a drivers' career. The expected market shares are those observed in the data; elasticities of demand are from model three in Table 6. Red dots indicate the return to effort for minority drivers and blue dots represent the corresponding results for nonminority drivers.

First, the impact of efforts on profits decreases as more information about drivers becomes already available. Second, the initial increase is higher for minority drivers. Two countervailing factors shape the disparity between minority and nonminority drivers: a higher variance of types in the population of minority drivers results in more uncertainty about individual types, and as a consequence, higher efforts. Although the expected profits in the first several rounds are smaller for minority drivers, which dampens the incentive to exert effort, the market shares increase over time, so that the latter effect is not particularly strong.

The incentives to exert effort are closely linked with the impact of a grade on future revenue. Table 7 shows the change in revenue following a grade from one to five. We take into account the elasticity of demand with respect to the number of reviews and quality. Only a grade of five has a positive impact. The grade of four leads to almost no change in revenue, and all lower grades result in negative and substantial changes. Minority drivers experience a more significant reaction to any grade because of the higher variance of types. They lose more as a result of low grades and experience a more significant benefit from a grade of five.

Figure 9: Incentives to exert effort



Note: Horizonal axis - number of reviews. Vertical axis - the impact of a unit of effort on future market shares. Red dots - minority drivers. Blue dots - nonminority drivers.

Table 7: Impact of a grade on revenue

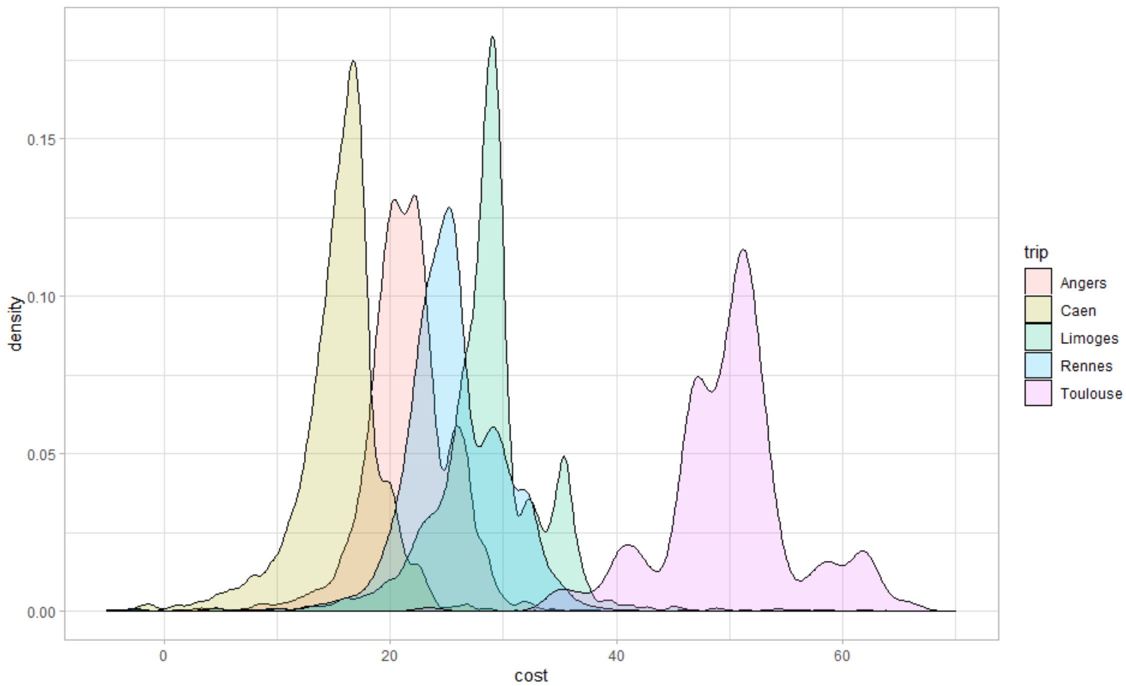
	1	2	3	4	5
Minority	-33.18%	-21.98%	-10.78%	-0.42%	11.63%
Nonminority	-45.76%	-29.75%	-11.74%	-5.28%	22.28%

Note: The figures show percentage changes of the next predicted revenue amount following a grade from 1 to 5. This impact arises from the the number of reviews and expected quality.

Marginal costs

Figure 10 shows the distribution of the recovered marginal costs on selected trips. These costs are related to trip length; long trips are associated with higher marginal costs than are shorter trips. The difference in marginal costs between minority and nonminority drivers (23.3 and 22.6, respectively) is 3.2%.

Figure 10: Marginal costs



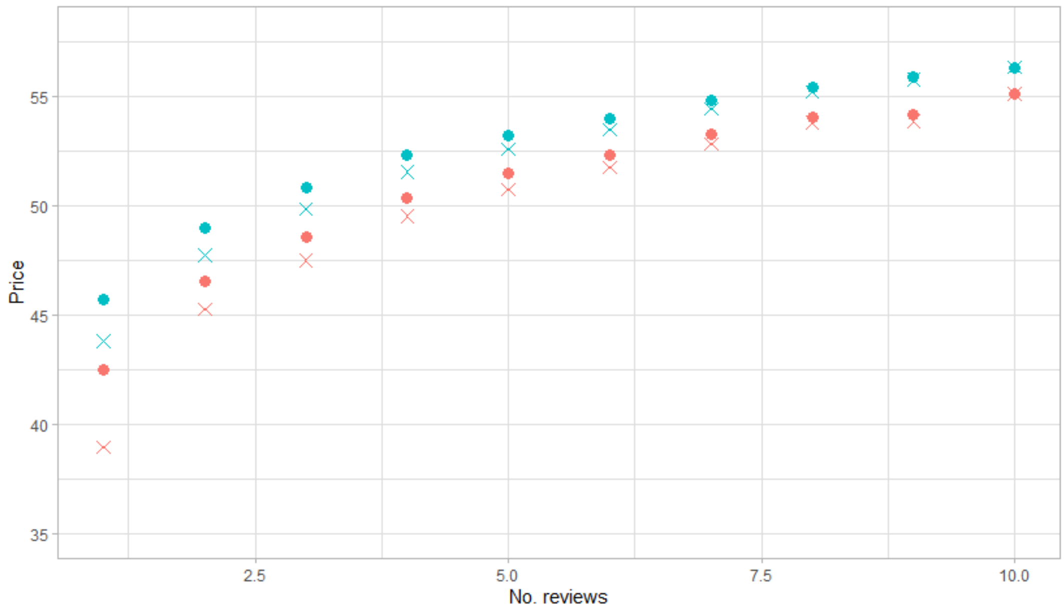
Note: Marginal costs given by equation 15 of drivers with more than ten reviews; selected trips

Pricing results

We are interested in how the incentive to invest in reputation translates into low introductory prices. Throughout this section, we will compare the prices set by drivers if they internalize the reputation-building incentive while setting prices and the prices set if drivers do not do so. Dynamic prices are solutions of equation **??**, while static prices satisfy equation 15.

Figure 11 compares static and dynamic prices of minority and nonminority drivers. We fix the marginal cost and all other driver-listing-specific characteristics, except for the number of reviews and the expected quality. The expected quality in each period equals the mean of expected qualities

Figure 11: Dynamic vs. static prices



Note: Horizontal axis - the number of reviews. Vertical axis- the optimal price. Minority drivers - red; nonminority - blue. Bullets- static profit-maximizing prices. Crosses- prices, resulting from internalizing reputation-building incentive.

of all drivers with the same set of characteristics.[26]

Optimal static prices (bullets in Figure 11) increase over time due to the positive elasticity of demand with respect to the number of reviews. Nonminority drivers receive, on average, the same reviews as the market expects. However, minority drivers experience an additional benefit from reputation because of the increase in posterior beliefs. Thus, even static prices increase more rapidly for minorities than for nonminorities.

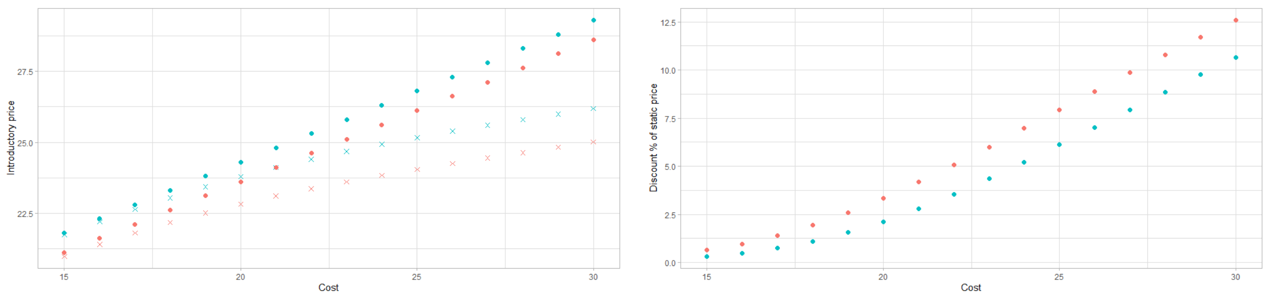
The prospect of higher profits motivates all drivers to act strategically and offer discounts. Dynamic prices start at lower levels (e.g., the first period's prices might be below costs), increase more rapidly, and converge to static prices at period ten. Note that under dynamic pricing, drivers sell seats faster during the first couple of periods. Minority drivers take into account the expected correction in the market belief about their quality. As a result, they offer larger discounts.

The change in prices from period to period also depends on marginal costs. Figure 12 repeats the exercise illustrated by Figure 11 but considers several levels of costs. We present the difference in introductory prices between static and dynamic pricing modes for different levels of costs. The

[26]The demand predicted for a ride with: a photo, the automatic acceptance feature, the maximum 2 passengers option, the ride occuring during the day, the time since the listing has been posted equal to the mean in the dataset, the notice equal to the mean in the dataset, seniority equal to the mean in the dataset, car price equals to the mean in the dataset, the ride occuring during a weekday on a non-strike day.

difference increases with marginal costs. Lowering introductory prices increases the probability of selling a seat and receiving a review. However, at lower levels of marginal costs, the prices are already relatively low in the static case. Thus, lowering them further has a proportionally smaller impact on increasing the chance of receiving a review. Furthermore, drivers with low marginal costs earn a significant markup even when they have only a few reviews. Thus, their incentive to invest in reputation is smaller.

Figure 12: Discount in introductory prices for various levels of marginal cost



Note: Static vs. dynamic introductory prices. Minority drivers - red, nonminority drivers - blue. The left panel compares dynamic prices (crossess) with static prices (bullets). The right panel shows the difference between static and dynamic as a percentage.

So far, we have focused on one driver with one set of characteristics. However, we have already observed that the incentive to offer a discount from static prices for drivers with few (or no) reviews depends on the level of marginal costs. It also depends on other drivers' characteristics.

Generally, the lower the initial market share is, the higher the relative increase in profits following a review. Therefore, to quantify the average discount in introductory prices, we set the parameters in the algorithm to match those of listings we have observed (thus, we assume that ride-specific parameters - photograph, automatic acceptance, weekday, etc. do not change as the driver receives reviews), and estimate marginal costs. We focus on the sample of markets used in demand estimation. Based on the recovered marginal costs, we compute introductory prices for a driver who follows Bertrand pricing and compare them with the observed dynamic prices.

We find that nonminority drivers reduce their prices by 4.08% on average, which is a significant investment in reputation. Minority drivers reduce prices by 8.03%; the larger discount is due to a higher increase in future market shares following an expected review. Consistently with the example in Figure 12, the difference is higher for drivers with higher marginal costs.

8 Counterfactual experiments

The structural model allows us to generate counterfactual experiments. We will analyze three alternative scenarios. First, we simulate market outcomes under the correct prior. In this scenario, passengers have correct beliefs about the expected quality of all minority drivers. Comparing the baseline scenario with this experiment allows us to calculate the cost to minority drivers of erroneous beliefs. Second, we study the market in which the gap between minority and nonminority drivers remains constant. In this case, the expected quality is always reduced by the size of the bias. This simulation highlights the difference between statistical discrimination (the baseline case) and taste-based discrimination. Finally, we evaluate a policy intervention proposed by Benjamin Edelman and Michael Luca (Edelman et al. (2017)) that makes the profiles of drivers ethnicity-blind. Table 8 summarizes the main results.[27]

Table 8: Summary of counterfactuals

	Δ quality	Δ efforts	Δ intro prices minority	Δ π minority	Δ π nonminority
Correct prior	2.9%	4.91%	3.91%	19.13%	-0.48%
Persistent bias	-4.95%	-14.28%	0.62%	-7.69%	0.11%
Debias yourself	10.03%	7.54%	13.34%	21.6%	-0.85%

Note: All values are percentage changes compared to the baseline. Column 1: average change in the expected quality of minority drivers on trips 1-15. Column 2: change in total efforts of minority drivers. Column 3: change in introductory price charged by minority drivers. Column 4: change in average profits of minority drivers over trips 1-15. Column 5: change in average profits of nonminority drivers over trips 1-15.

Cost of the incorrect prior

This exercise aims to quantify the cost of erroneous beliefs. Under this scenario, minority drivers will be evaluated ex ante following their true quality, as revealed by the grades they obtain ex post.

This change spurs several reactions. First, minority drivers will be perceived by the market as being of higher quality. They will be able to raise prices and exert more effort, so their quality will increase further. Nonminority drivers will react to this by reoptimizing their prices. Finally, passengers

[27] In each of the scenarios, we characterize a new equilibrium described by definition **??**. Each of the proposed counterfactuals involves changes in passenger decisions, which leads to new optimal prices and efforts by both minority and nonminority drivers, which again lead to a different set of passenger decisions. Thus, we are looking for new vectors of purchasing decisions, pricing, and efforts such that none of the parties can gain by deviating.

in the counterfactual markets will choose between minority drivers, whom they now perceive to be of higher quality, but who now charge higher prices, nonminority drivers with new levels of prices, and the outside option, that is unchanged.

This scenario assumes that the belief about the expected type of a minority driver with no reviews improves from 4.1 to 4.49 (a change from the 7.5th percentile to the 50th). The process of updating beliefs about individual drivers' quality proceeds the same way as before. As a consequence, throughout the first 15 periods, the average perception of the expected quality of minority drivers increases by 2.9%. Moreover, the amount of effort increases by 4.91%, further boosting quality. The optimal level of effort is particularly susceptible to changes in expected profits in the first period; hence, the sizable change.

The higher expected quality allows minority drivers to increase introductory prices. The incentive to reduce the price to hasten belief correction disappears. Introductory prices rise by 3.91%. A higher expected quality and the change in prices have a substantial effect on expected profits that increase by 19.13%.

In other words, 19.13% of profits is the price minority drivers have to pay for incorrect beliefs held by passengers. Finally, the profits of nonminority drivers decline by 0.48%. Most of the change in substitution is with respect to the outside option.

Persistent bias

Suppose that the bias against minority drivers is not subject to change. Each driver can have her individual reputation, but minority drivers are always considered to be worse, regardless of how many reviews they have. Let the size of this bias be given by the extent to which the expected belief about the type of minority drivers differs from the mean type revealed by the grades (4.1 vs. 4.49).

If there is no possibility of mitigating discrimination, minority drivers always achieve lower profits, and their incentives to exert effort vanish. The exerted efforts decline by 14.28%, further depressing the quality of service provided by minority drivers.

Interestingly, in this case, minority drivers will charge higher introductory prices. This is so because the expected quality of service of a minority driver with no reviews is the same as in the baseline case, but the incentive to reduce the price to receive more reviews is lower. The average profit throughout the first 15 periods is lower by 7.69%. There is little substitution away from minority drivers to nonminority drivers, whose profits increase by 0.11%.

Ethnicity-blind profiles

The publication of the studies by Edelman and Luca (2014) and Edelman et al. (2017) that were among the first to document racial bias on sharing economy platforms spurred a heated discussion about ways to address the problem. One of the proposals was to change the way platforms displayed ethnicity or gender-related information. To this end, the authors of the above papers developed a web browser plugin called *Debias Yourself*[28] that removed names and photos of hosts on Airbnb.

Airbnb itself started addressing the issue of racial bias by changing the way profiles were presented. In 2016, the listing page (the page displayed after a search query) stopped showing names and photos of hosts. Only information specific to the listing became available. To view host-specific information, a potential guest had to click on the listing.[29]

For this experiment, let us suppose that a passenger does not know whether the driver with whom she is planning on booking a trip is a minority. It is also impossible to deduce that from other observables. Therefore, the passenger forms an expectation based on the distribution of drivers in a given market. The share of minority drivers differs depending on the route. The highest ratio of minority to nonminority drivers is on the route from Lyon to Paris (16%), and the lowest is on the Rennes-Paris connection (7%).[30]

As a result, the market perceives minority drivers to have an expected quality that is higher by 10.03%. Drivers will react to this policy by reoptimizing effort levels and their pricing strategies. Now, both minority and nonminority drivers have an incentive to reduce their introductory prices because reviews improve the beliefs about quality for everyone. However, for minority drivers, this incentive is lower than in the baseline case. From a static perspective, minorities should increase prices immediately because their quality is now believed to be higher. Considering both effects, we observe that introductory prices set by minority drivers increase by 13.34%. By the same logic, nonminority drivers reduce their prices.

The increase in the expected quality and the rise in prices result in higher profits of minority drivers; the latter increase by 21.6%. Nonminority drivers earn slightly lower profits, a reduction of 0.85%. The change in expected profits of nonminority drivers is more substantial for drivers with high marginal costs. In Figure 13, we show the change in percentage terms of the discounted sum of

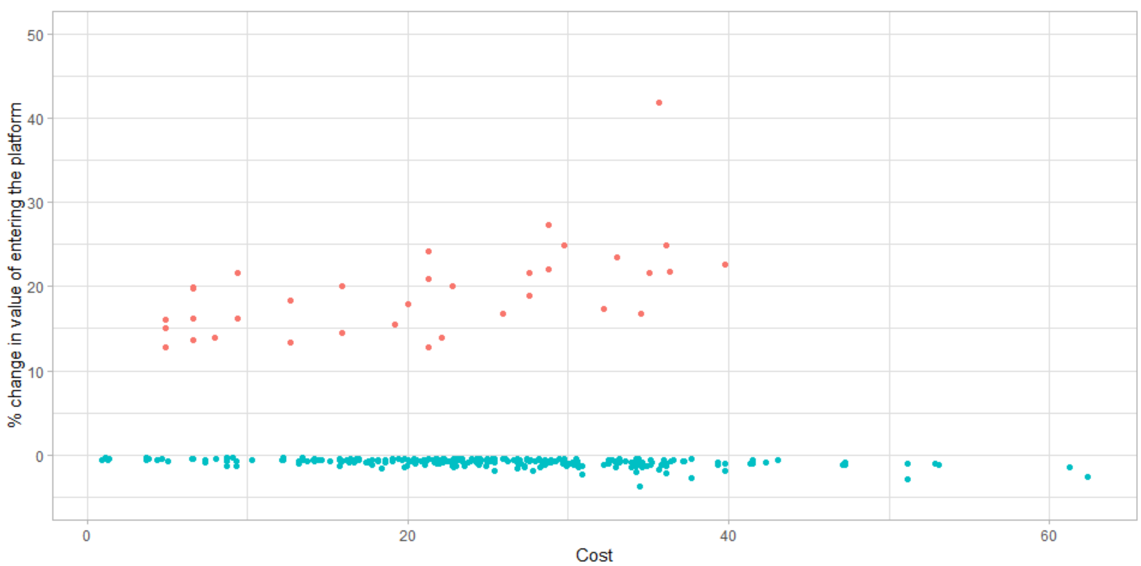
[28] http://www.debiasyourself.org/index.html

[29] See https://www.cnbc.com/2017/04/07/airbnb-experimenting-with-site-design-to-fight-discrimination.html for details.

[30] We also assume that reviews do not reveal ethnicity.

profits. This experiment reveals that if drivers are heterogeneous in unobservables, imposing a veil of ignorance on some observables might have unintended consequences.

Figure 13: Percentage change of the discounted sum of profits.



Note: Minority drivers - red. Nonminority drivers - blue. Results are for randomly selected 500 drivers. Horizontal axis- marginal cost. Vetical axis- change in the discounted sum of profits earned in the counterfactual scenario.

In this paper, we do not model entry into the market. However, given the changes in expected sums of profits in all three counterfactual scenarios, we should expect a change in the composition of drivers. Minority drivers have stronger incentives to join the platform when their expected quality is believed to be higher and under ethnicity-blind profiles. They would be less likely to enter the market when they face a persistent bias. The incentives to enter the market for nonminority drivers are changing in precisely the opposite direction.

9 Conclusions

Online discrimination against minorities has been documented in many prominent marketplaces. In this paper, we show that in the context of BlaBlaCar, a significant part of discrimination arises due to incorrect and overly pessimistic prior beliefs about the quality of service offered by minority drivers. These beliefs are altered with reviews. The initial gap of approximately 12% in revenue declines as minority drivers accrue reputation. The revenue differential for experienced drivers is statistically insignificant. The improvement in the performance of minority drivers is due to a causal effect of reviews, as we show using a difference-in-differences analysis.

This paper provides evidence that minority drivers use the reputation system to their benefit. They increase their levels of effort to receive high grades and set low introductory prices to build up their reputations faster. In the context of BlaBlaCar, the online reputation system allows mitigating ethnic discrimination. However, this is a costly fight for minority drivers. They have to persevere through an initial period of low economic outcomes and invest in their reputation. To calculate the cost of incorrect beliefs, we perform a counterfactual experiment. We simulate market outcomes in a scenario in which the initial beliefs about the quality of service of minority drivers are correct. Over the first 15 rides, we observe an increase in profits by 19%, which is the true cost of incorrect beliefs.

We propose a model of career concerns that represents a novel approach to studying the incentives of sellers in online markets. A reputation system creates an intertemporal externality. Reports of past performance can reveal some demand-relevant and seller-specific information and, as a result, boost or hurt future outcomes. However, the seeming randomness of reviews makes the task of extracting information out of grades difficult; this is why we need a model. We indeed show that reviews exhibit random components. Nevertheless, the model we propose allows us to separate the random element from the information that the market can use to update beliefs about the expected future quality. The ratings of minority drivers are on average higher than the market expectation. Thus, such drivers' quality of service is typically believed to be higher a posteriori than a priori.

The platform itself does not create prejudice against minorities. However, platform design can both mitigate discriminatory behavior and exacerbate it. BlaBlaCar provides information that reveals ethnicity of drivers, which allows passengers to discriminate based on it. The platform also equips minority drivers with tools to counter discrimination. The entire history of reviews is available on profiles of drivers, which helps inform future passengers. Drivers can also influence the speed of beliefs' updating by offering discounted prices. Thus, BlaBlaCar's online infrastructure enables its users to alter their incorrect priors.

This paper contributes to a long-standing discussion of the sources of discrimination. In our context, discrimination is to a large extent due to incomplete information. Passengers on BlaBlaCar are willing to change their beliefs when they are presented with an additional review. This result has clear policy implications: the provision of information is an effective way of tackling discrimination, at least in this market.

References

Abadie, A. and Imbens, G. W. (2016). Matching on the estimated propensity score. *Econometrica*, 84(2):781–807.

Abrahao, B., Parigi, P., Gupta, A., and Cook, K. S. (2017). Reputation offsets trust judgments based on social biases among airbnb users. *Proceedings of the National Academy of Sciences*, 114(37):9848–9853.

Agrawal, A., Lacetera, N., and Lyons, E. (2016). Does standardized information in online markets disproportionately benefit job applicants from less developed countries? *Journal of international Economics*, 103:1–12.

Alesina, A., Michalopoulos, S., and Papaioannou, E. (2016). Ethnic inequality. *Journal of Political Economy*, 124(2):428–488.

Altonji, J. G. and Pierret, C. R. (2001). Employer learning and statistical discrimination. *The Quarterly Journal of Economics*, 116(1):313–350.

Arrow, K. J. (1973). The theory of discrimination, discrimination in labor markets. *Achenfelter, A. Ress (eds.), Princeton–New Jersey*.

Banerjee, A. and Munshi, K. (2004). How efficiently is capital allocated? evidence from the knitted garment industry in tirupur. *The Review of Economic Studies*, 71(1):19–42.

Bartoš, V., Bauer, M., Chytilová, J., and Matějka, F. (2016). Attention discrimination: Theory and field experiments with monitoring information acquisition. *American Economic Review*, 106(6):1437–75.

Becker, G. (1971). The economics of discrimination. *University of Chicago Press Economics Books*.

Bohren, J. A., Haggag, K., Imas, A., and Pope, D. G. (2019a). Inaccurate statistical discrimination. Technical report, National Bureau of Economic Research.

Bohren, J. A., Imas, A., and Rosenberg, M. (2019b). The dynamics of discrimination: Theory and evidence. *American Economic Review*, 109(10):3395–3436.

Carol, S., Eich, D., Keller, M., Steiner, F., and Storz, K. (2019). Who can ride along? discrimination in a german carpooling market. *Population, Space and Place*.

Chiappori, P.-A., Salanie, B., and Valentin, J. (1999). Early starters versus late beginners. *Journal of Political Economy*, 107(4):731–760.

Coate, S. and Loury, G. C. (1993). Will Affirmative-Action Policies Eliminate Negative Stereotypes ? *The American Economic Review*, 83(5):1220–1240.

Conlon, C. and Gortmaker, J. (2019). Best practices for differentiated products demand estimation with pyblp. Technical report, Working paper. url: https://chrisconlon. github. io/site/pyblp. pdf.

Cui, R., Li, J., and Zhang, D. J. (2019). Reducing discrimination with reviews in the sharing economy: Evidence from field experiments on airbnb. *Management Science*.

Edelman, B., Luca, M., and Svirsky, D. (2017). Racial Discrimination in the Sharing Economy: Evidence from a Field Experiment. *American Economic Journal: Applied Economics*, 9(2):1–22.

Edelman, B. G. and Luca, M. (2014). Digital Discrimination: The Case of Airbnb.com. *SSRN Electronic Journal*.

Farajallah, M., Hammond, R. G., and Pénard, T. (2019). What drives pricing behavior in peer-to-peer markets? evidence from the carsharing platform blablacar. *Information Economics and Policy*.

Gandhi, A., Lu, Z., and Shi, X. (2013). Estimating demand for differentiated products with error in market shares. Technical report, cemmap working paper, Centre for Microdata Methods and Practice.

Ge, Y., Knittel, C., MacKenzie, D., and Zoepf, S. (2016). Racial and Gender Discrimination in Transportation Network Companies. *NBER Working Paper Series*, (22776):1–38.

Glover, D., Pallais, A., and Pariente, W. (2017). Discrimination as a self-fulfilling prophecy: Evidence from french grocery stores. *The Quarterly Journal of Economics*, 132(3):1219–1260.

Heckman, J. J., Ichimura, H., and Todd, P. E. (1997). Matching As An Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme. *The Review of Economic Studies*.

Hjort, J. (2014). Ethnic divisions and production in firms. *The Quarterly Journal of Economics*, 129(4):1899–1946.

Holmström, B. (1999). Managerial incentive problems: A dynamic perspective. *The Review of Economic Studies*, 66(1):169–182.

Iacus, S. M., King, G., and Porro, G. (2009). Cem : Software for coarsened exact matching. *Journal of Statistical Software*.

Kakar, V., Voelz, J., Wu, J., and Franco, J. (2018). The visible host: Does race guide airbnb rental rates in san francisco? *Journal of Housing Economics*, 40:25–40.

Kuznets, S. (1955). Economic growth and income inequality. *The American economic review*, 45(1):1–28.

Laouenan, M. and Rathelot, R. (2017). Ethnic Discrimination on an Online Marketplace of Vacation Rentals. *working paper*.

Mazzella, F. and Sundararajan, A. (2016). Entering the Trust Age. Technical report.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. *Frontiers in Econometrics*.

Pallais, A. (2014). Inefficient hiring in entry-level labor markets. *American Economic Review*, 104(11):3565–99.

Phelps, E. S. (1972). The Statistical theory of Racism and Sexism. *American Economic Review*, 62(4):659–661.

Reynaert, M. and Verboven, F. (2014). Improving the performance of random coefficients demand models: the role of optimal instruments. *Journal of Econometrics*, 179(1):83–98.

Rosenbaum, P. R. and Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*.

Sant'Anna, P. H. and Zhao, J. (2020). Doubly robust difference-in-differences estimators. *Journal of Econometrics*, 219(1):101–122.

Sarsons, H. (2017). Interpreting Signals in the Labor Market: Evidence from Medical Referrals. *Job Market Paper*.

Tadelis, S. (2016). Reputation and feedback systems in online platform markets. *Annual Review of Economics*, 8:321–340.

Tjaden, J. D., Schwemmer, C., and Khadjavi, M. (2018). Ride with me -ethnic discrimination, social markets, and the sharing economy. *European Sociological Review*, 34(4):418–432.

Zervas, G., Proserpio, D., and Byers, J. (2015). A First Look at Online Reputation on Airbnb, Where Every Stay is Above Average. *Working paper*, pages 1–22.

A Navigation on Blablacar.fr

First, users type in the origin, destination and date of the ride they are seeking. They then see a list of rides meeting their request (Figure 14). They may then click on specific postings to have more details about the ride (Figure 15). Finally they may either see the profile of the driver (Figure 16) or proceed directly to payment. BlaBlaCar service fees are a function of the price posted by the driver. The fees and their evolution over time are shown on Figure 17.

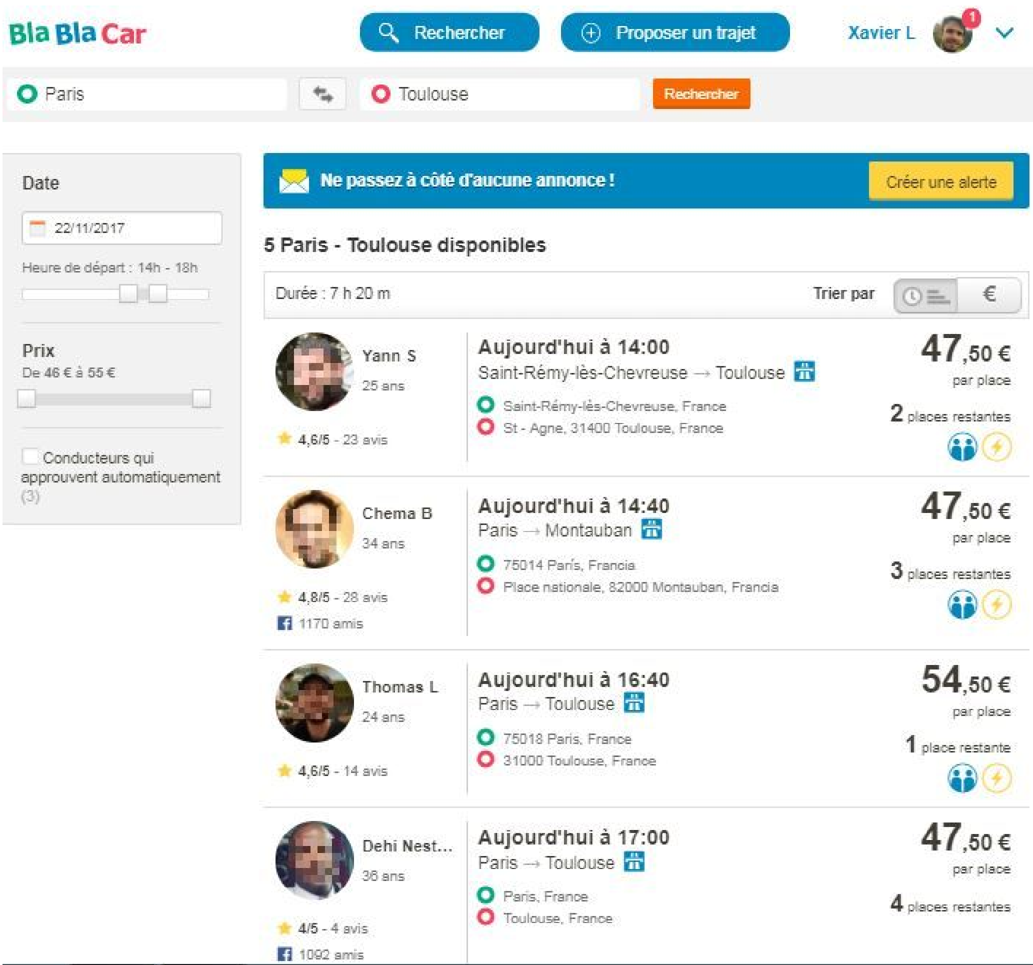


Figure 14: Listing offered on a given route

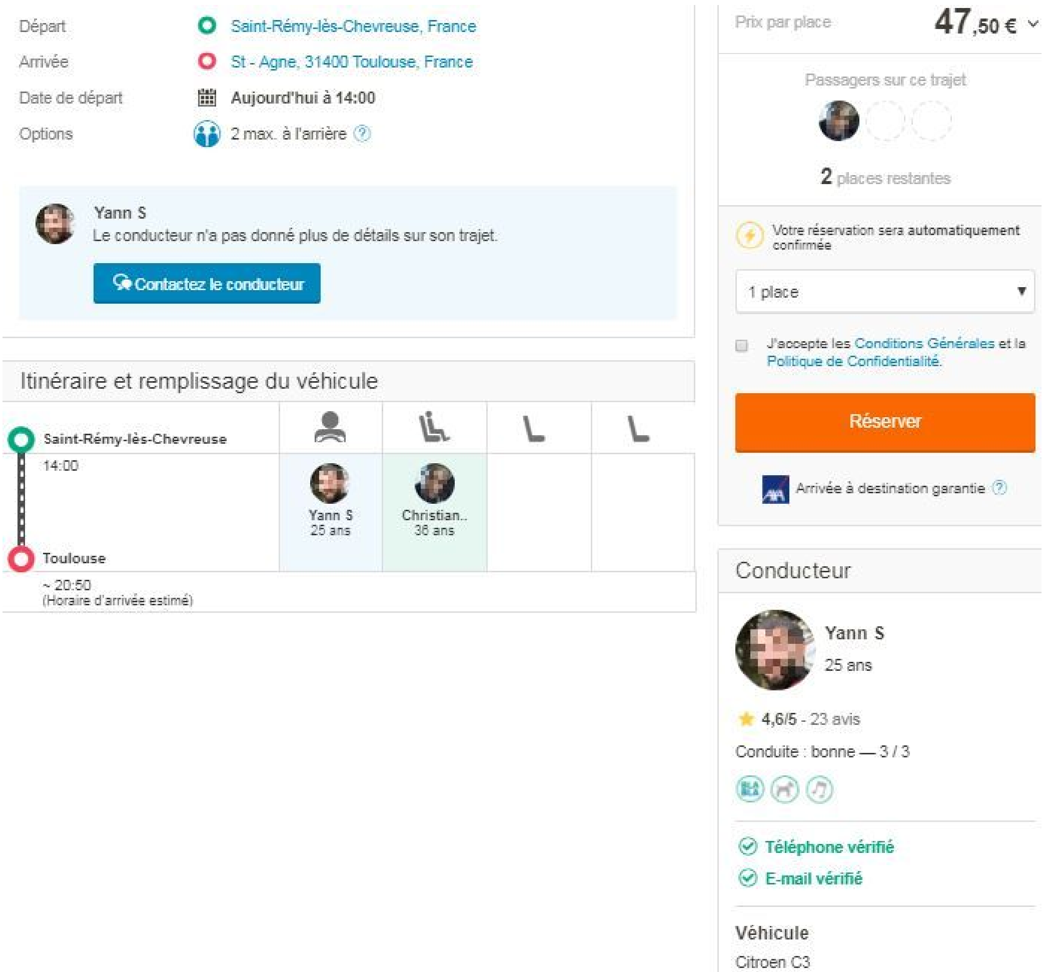


Figure 15: Details of a posting

figs/profile_sample.jpg

Figure 16: A driver's profile

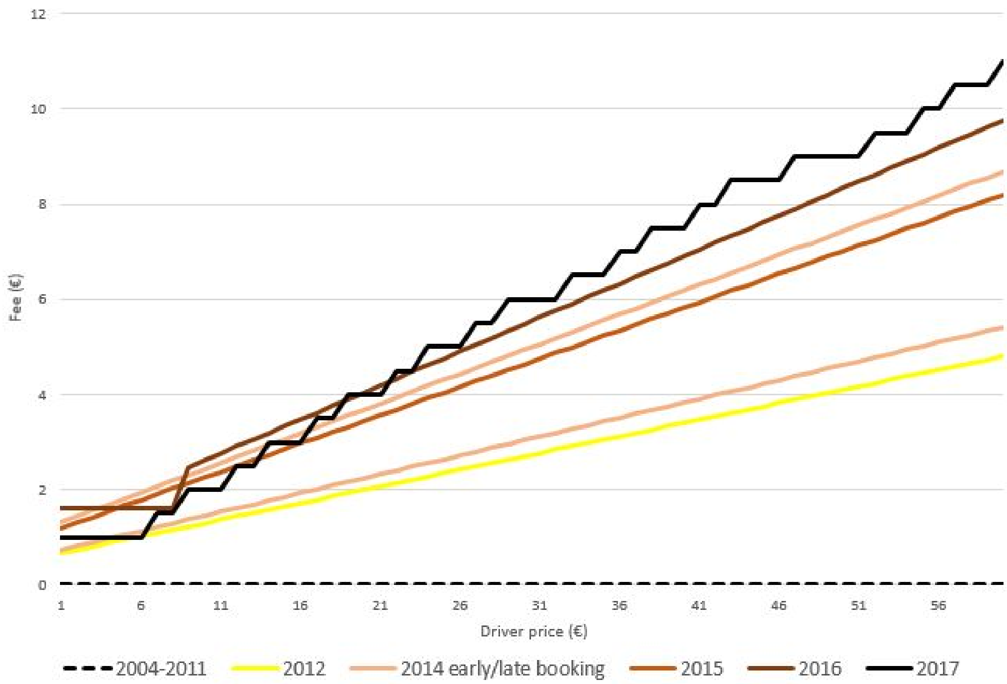


Figure 17: Evolution of service fees on BlaBlaCar over time

B Changes in the BlaBlaCar reputation system

In our study of the evolution of ratings, we have abstracted from the potential changes in the design of the reputation system of BlaBlaCar. Some drivers in our sample have been BlaBlaCar users since December 2008, and others joined only a few days before our crawler observes their listing. These drivers may have operated under different market characteristics. See Figure 18 for the evolution of the average rating over time. Until the end of 2013, ratings were either 1 or 5. In early 2014, these binary ratings were translated to the current 5-star system. Later, in February 2016, the wording of the ratings was changed: *excellent* became *tres bien* and *extraordinaire* became *parfait*. The impact of this change on the average rating is clear. People are more likely to call a ride *parfait* than they were to call it *extraordinaire*. Finally, these changes influenced the informativeness of the reputation system; see Figure 19. The dotted black line shows HHI (which is a measure of dispersion and, hence, the informativeness of the classifiers): the smaller the HHI is, the more informative the classifier. The ratings in the period 2014-2016 were the most informative. Dark green, green, orange, pink, and red represent the shares of 5s, 4s, 3s, 2s and 1s, respectively. Initially, there is a considerable noise because we have very few observations: fewer than 100 per month before October 2009 and more than 30.000 per month starting in 2017.

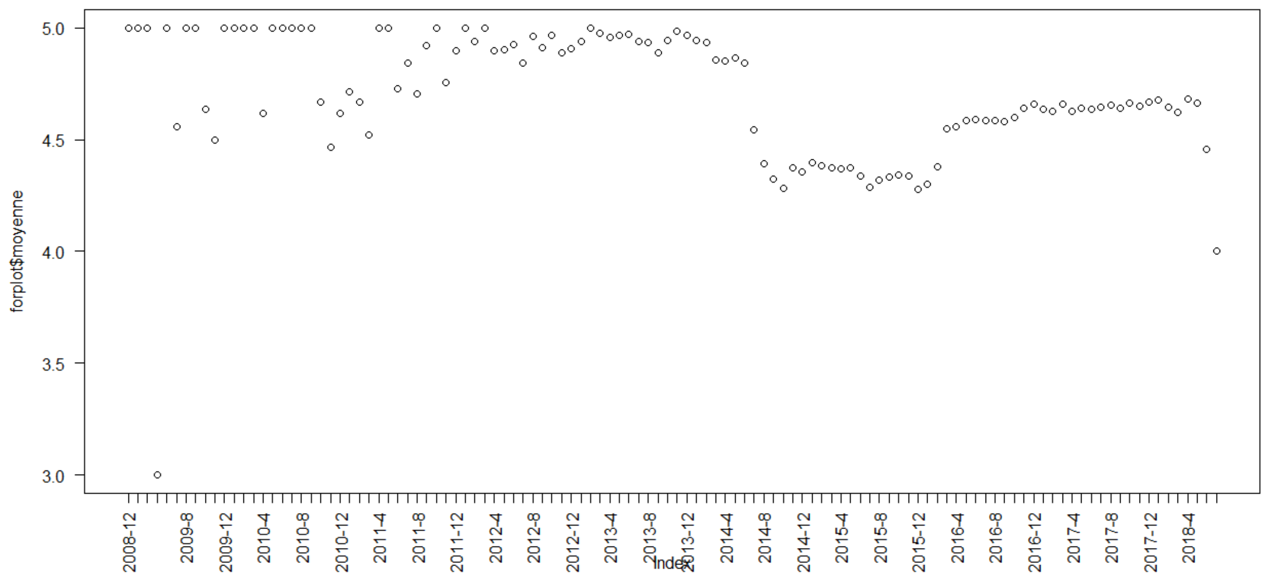


Figure 18: Average rating for drivers with more than 30 reviews

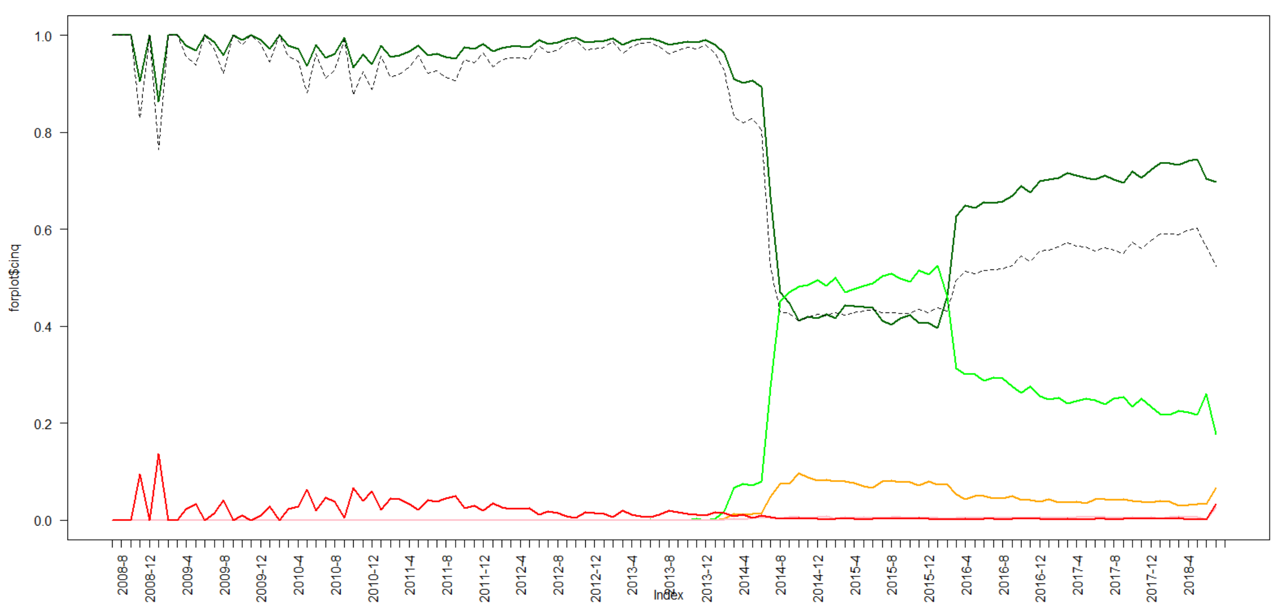


Figure 19: Informativeness of the reputation system and share of grades received. Dark green=5, light green=4, orange=3, pink=2, red=1.

These changes are important because they affected the ratings that we study, but they also show how important the design of the review system is. One may be concerned that some of the decline in effort that we characterized could be due to changes in the reputation system. In a sample restricted to drivers who joined after all the changes in the reputation system were made, we can reproduce the same patterns of behavior; however, we lose a considerable number of observations. Thus, we argue that the evolution of ratings throughout the career of a driver on BlaBlaCar is due to the economic logic of career concerns rather than exogenous changes in the reputation system.

C Classification method for gender and ethnicity

Driver-specific characteristics are key determinants in our model. Hence, the drivers' type must be identified as accurately as possible. Specifically, gender and ethnicity are critical to our analysis. To identify these characteristics, both prospective riders and the econometrician consider two relevant sources of information: the first name and the profile picture. We use both information to infer gender and ethnicity.

C.1 Classification of gender

As a first source of information, we use the name of the driver. We match our dataset of driver names with those of various sources relating first names with ethnicity. The French Government repository of names (www.data.gouv.fr/fr/datasets/liste-de-prenoms) constitutes our main source of information. We complement it with data from other sources.[31] This data enables us to identify the gender of almost 80% of drivers, along with 3% unisex names.

We then use facial recognition to identify gender whenever a picture is available. This process also enable us to identify 80 % of the dataset. By combining these two processes, we can directly identify gender for 95% of the dataset.

Further, we use facial recognition to enrich and correct our name database. Rare or misspelled names (either because the driver registered under a nickname or because of translation variations if the name is not originally French) can be re-classified. This process can identify the gender of some drivers whose names are not listed in our inventories and who do not have a picture (or for pictures where gender is not easily identified) because other drivers with the same name may have posted identifiable pictures. This method brings the precision of our gender identification as high as 99%. Figure 20 summarizes our identification process.

[31] www.signification-prenom.net, www.madame.lefigaro.fr/prenoms/origine

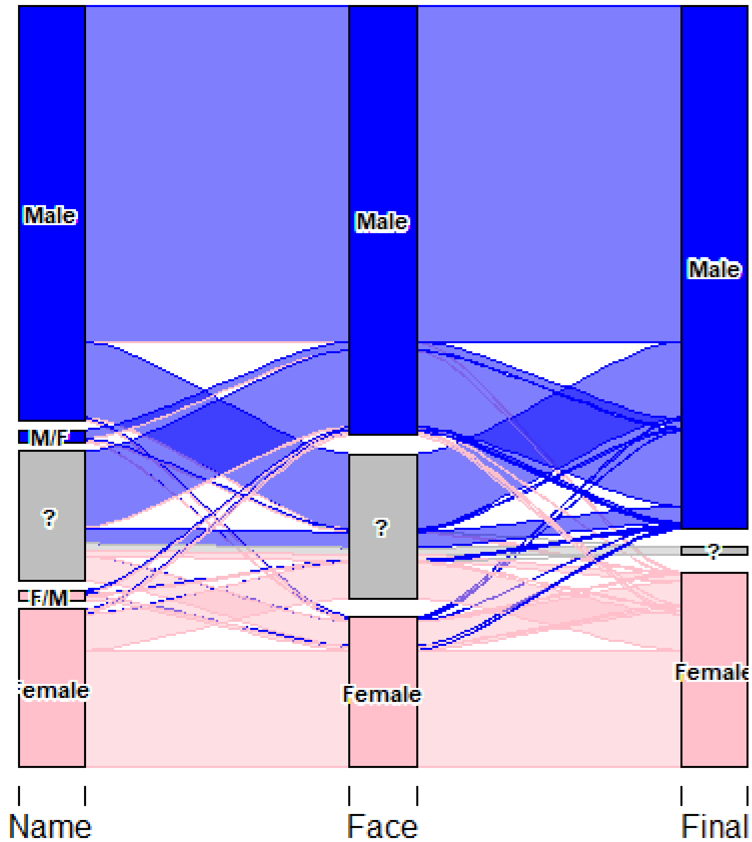


Figure 20: Classification process for gender: by name (left), by facial recognition (center) and final classification (right)

C.2 Classification of ethnicity

Our methodology for the identification of ethnicity follows the same steps and uses the same sources as those for gender classification. First, we collect the origins of names from the data sources mentioned above. This provides the ethnicity of approximately 81% of our sample. However, names might not be a perfect indicator of ethnicity. Indeed, many visible minorities have a French name for various historical reasons or because they have foreign origins but were born in France. In that case, a simple name analysis would classify them as non-minorities while they might belong to a minority on the basis of their skin color.

Hence, we use facial recognition to identify ethnicity whenever a picture is available. The algorithm proposes an ethnicity for 80 % of the dataset. However, only "white", "black", "Asian" , and "Latino" ethnicities are proposed. People of Arabic origin are classified as "white". Hence, facial recognition is useful only to classify drivers more accurately between african origin, and majority or arabic origin.

We also use facial recognition to enrich and correct our name repository and to better identify ethnicity. Overall, facial recognition reclassifies 2.5% of drivers with a French name and 5% of drivers with Arabic names (predominantly Muslim names) into Sub-Saharan ethnicity. Including facial recognition increases the sample size for minorities from 11% to 14% of our sample. Figure 21 summarizes our identification process.

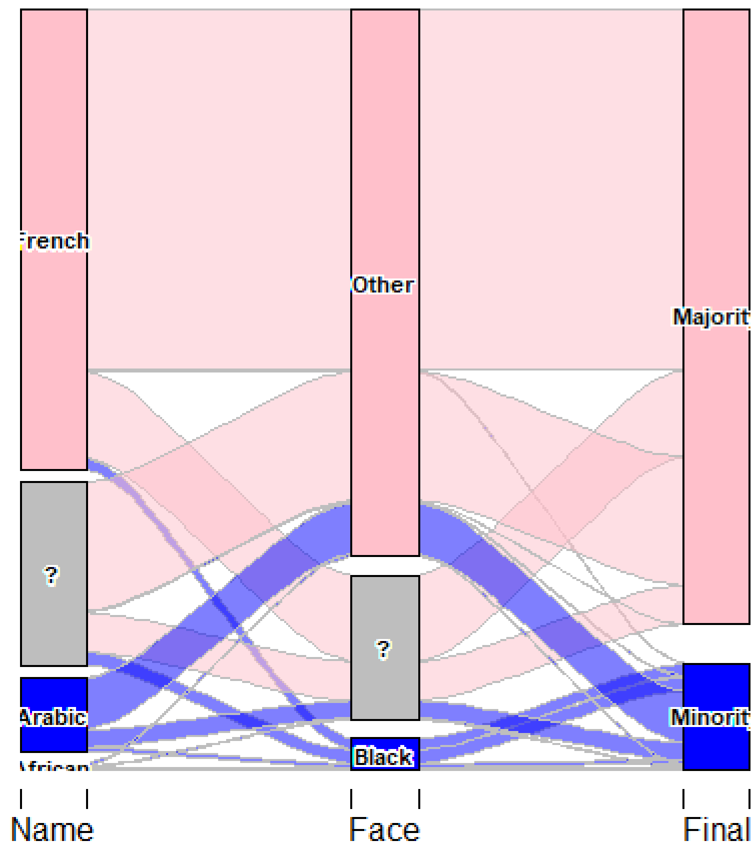


Figure 21: Classification process for ethnicity: by name analysis (left), by facial recognition (center) and final classification (right)

D Ratings as a measure of passenger satisfaction

The body of the paper analyses the effect of reputation on the sole basis of ratings. It assumes that ratings have enough informational content to allow passengers to form a belief about the quality of a driver.

In this Appendix, we show that ratings are indeed likely to be a good summary of passengers' experience. To do so, we analyze whether good reviews (i.e. reviews with a high rating) are more likely to be associated with a written comment that has a positive connotation than bad reviews. For that purpose, we use the Cloud Natural Language processing tools of Google, a tool that uses machine learning to reveal the structure and meaning of text. We are particularly interested in the sentiment of the review, with a measure between -1 (very negative) and 1 (very positive).

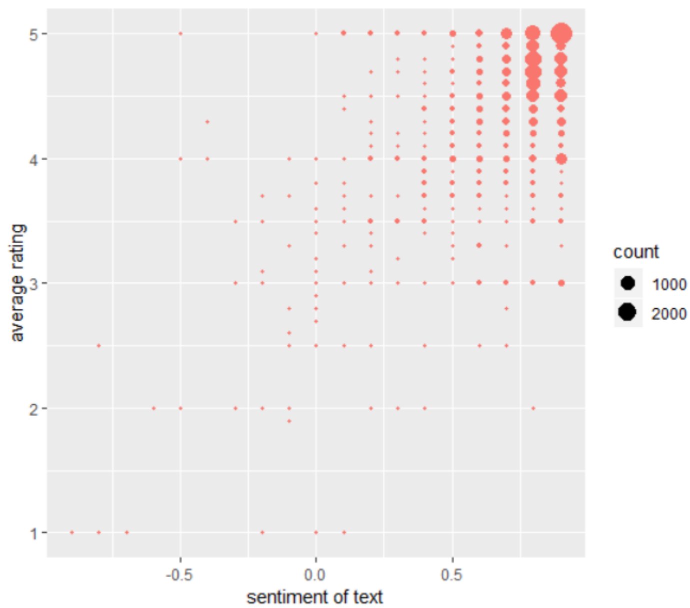


Figure 22: Textual analysis of the comments (18000 randomly selected drivers). Average rating and average sentiment of written comment are highly correlated.

The correlation between the grade given, and the sentiment of the text of the review very high, as is suggested by Figure 22. We therefore conclude that ratings are a satisfactory a measure of performance for the purpose of the present paper.

E Oversampling of minorities for short-notice rides

Due to our scraping method, it cannot be excluded that our sample provides a slightly biased representation of listings. Indeed, the program takes snapshots of listings displayed on the website at a given point time. However, rides that are already full are no longer displayed on the platform. This means our data collection may undersample the particularly attractive rides that would sell out very fast, or those corresponding to times when demand is much higher than supply. This wouldn't be an issue if both minorities and non-minorities were affected the same way by this sampling bias. However, as we show in this paper the minority status does impact the attractiveness of a given listing. Therefore, minorities who may be perceived as posting less attractive rides remain longer on display and may therefore be over-represented in our sample. Therefore, our minority gap estimates should be understood as lower bounds. Indeed, minorities are compared to a pool constituted of non-minorities that are not so good as to have sold out their seats extremely fast. Table 23 shows that minority drivers represent a specially high share of rides that are posted on a short notice, a possible sign that non-minority drivers have sold their seats faster. For trips posted with more notice, we believe our sample is indeed representative of the actual participants on blablacar. Indeed, most of the rides –either from minorities or not – still have more than one empty seat, which means that most listings and indeed collected. In fact, Blablacar informs drivers that most passengers book rides only a few days in advance.

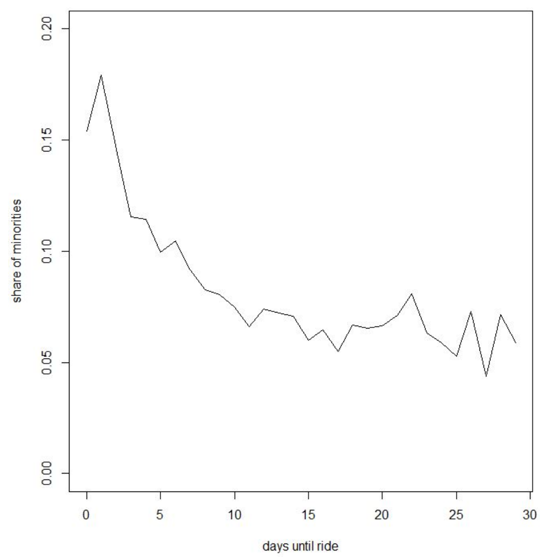


Figure 23: Share of minorities in sample as a function of number of days between posting and departure

This is true despite the fact minorities tend to allow for automatic confirmation more frequently than non-minorities (18% of drivers with automatic confirmation are minorities, while they represent only 12% of the drivers with manual confirmation).

F Definition of variables

Sources of supplementary data

- Databaset of names constructed based on: French government statistics www.data.gouv.fr/fr/datasets/liste-de-prenoms and supplemented with www.signification-prenom.net , www.madame.lefigaro.fr/prenoms/ori
- Car prices on eBay Germany: www.kaggle.com/orgesleka/used-cars-database
- Fuel consumtpion of cars: French environment and energy management agency- ADEME

name of a variable	description
price	price set by the driver in EUR; has to be lower than maximum price: 0.082 per km
age	age of the driver in years
reviews	number of reviews received by the driver
male	gender defined based on photo recognition and name
minority	takes the value of one when the driver is of Arabic or African origin, and zero otherwise; defined based on photo recognition and name (see. Lambin& Palikot (2019)) for details)
picture	takes the value of one when driver added a picture, and zero otherwise
talkative	categorical variable (bla, blabla, blablabla) indicating how talkative the driver is
bio	number of words in driver's description
ride description	number of words in ride's description
reputation	mean of grades received by the driver
published rides	number of rides ever published by the driver
number of clicks	number of clicks a given listing has received; clicking is necessary for booking a ride but not sufficient; measured at the moment of data collection
sold seats	number of seats already sold; measured at the moment of data collection
revenue	sold seats multiplied by price
posts per month	mean number of listings posted by the driver since she joined the platform
seniority	number of months since the driver joined the platform
competition	number of listings available on the same day on the same route
median revenue	mean of median revenues in cities of departure and arrival; source: INSEE
public transport	travelling time by public transport on the route at listings' departure time; source: Google API
train strike	SNCF official strike implicating a given route
value of car	price of a comparable car model in thousands of EUR; when a model of a car is not available mean price of a brand; source: ebay (scrapped data)
fuel consumption	mean fuel consumption of a model of a car; when model of a car is not available mean consumption of a brand; source: ADEME
length (km)	distance in km between cities of departure and arrival; souce: Google API
lengh (hours)	estimated driving time by a car on a given route and time; source: Google API
hours until departure	number of hours between data collection and a ride departure
posted since	number of hours between the posting of the listing and data collection
automatic acceptance	takes the value of one if booking requests are automatically accepted and zero if the driver chose to accept/reject requests manually
to fuel price	average price of a litre of diesel in a city of arrival in cents
from fuel price	average price of a litre of diesel in a city of departure in cents
toll viamich	total toll costs on a given route in EUR; source: https://www.viamichelin.com/
travel costs	mean of fuel costs multiplied by fuel consumption plus toll fees
weekday	takes a value of 1 on weekdays and zero on weekends
pets	takes a value of 1 if the driver accepts pets and zero otherwise
music	takes a value of 1 if the driver listens to music in the car and zero otherwise
smoke	takes a value of 1 if the driver accepts smoking in the car and zero otherwise
detour	categorical variable: 1 if no detour, 2 if some detour (up to 15 min), and 3 if more than 15 minutes detour
luggage	categorical variable: 1 if no luggage, 2 if small bags, 3 if big bags are allowes

Table 9: Defintion of main variables

- City specific population, median income, index of crime, and a share of foreign born residents- French statistics office INSEE.

G Reduced-form evidence

Market-specific effects and other observed characteristics of drivers could explain these differences. We will now control for all variables available in our dataset. Throughout the paper, subscript i refers to drivers. We estimate the following model:

$$y_{itr} = \alpha + X_{it}\beta + Z_i\gamma + \tau_t + \xi_r + \epsilon_{itr}, \tag{17}$$

where t represents time, r corresponds to a route; y_{itr} is the variable of interest (i.e., the number of clicks or sold seats or the revenue), α is an intercept, X_{it} is a vector of time-varying explanatory variables , Z_i are time-invariant explanatory variables, τ_t denotes time effects, ξ_r is an effect specific to a route (a pair of cities), and ϵ_{itr} is the error term.

Table 10 presents the estimation results. The dependent variable in the first regression is the number of clicks; it is the number of sold seats in the second regression and revenue in the last one.

First, minority status has a negative coefficient and is highly statistically significant for all measures of economic performance. Second, the number of reviews has a positive impact and is highly statistically significant in all regressions. Note that increasing the number of reviews benefits both minority and nonminority drivers. The negative coefficients associated with the squared number of reviews suggest decreasing returns to accumulating reviews. Finally, younger drivers with rides that include extended descriptions experience better economic outcomes. After we control for the number of reviews, seniority on the platform has a negative coefficient.[32]

Reputation effect: When a driver has no reviews, passengers have to rely entirely on socioeconomic characteristics (age, gender, and ethnicity) to form beliefs about the expected quality of service. As the driver uses the platform, reviews left by past passengers become available on her profile and reveal individual information about the driver. As a consequence, the role of socioeconomic characteristics diminishes as the driver collects reviews.

If initial discrimination is due to incorrect beliefs about the expected quality of service provided by minority drivers, the intergroup disparity in economic performance will decline as individual infor-

[32]In Appendix H, we control for price in a regression that uses the number of sold seats as the dependent variable; we also instrument prices with cost shifters to address the endogeneity of price and quantity.

Table 10: Output measures regressed on driver and ride characteristics.

	Dependent variable:		
	Number of clicks	Sold seats	Revenue
Minority	−0.444*** (0.082)	−0.017*** (0.003)	−0.588*** (0.079)
Number of reviews	0.033*** (0.001)	0.002*** (0.0001)	0.041*** (0.001)
(Number of reviews)^2	−0.0001*** (<0.0001)	−0.00000*** (<0.0001)	−0.0001*** (<0.0001)
Male	−1.400*** (0.064)	0.002 (0.002)	−0.094 (0.061)
Driver age	−0.058*** (0.002)	−0.001*** (0.0001)	−0.022*** (0.002)
Posts per month	−0.557*** (0.020)	−0.010*** (0.001)	−0.201*** (0.019)
Bio (number of words)	0.001 (0.003)	0.0001 (0.0001)	0.006** (0.003)
Car value	0.006 (0.006)	−0.0001 (0.0002)	−0.010* (0.005)
Seniority (number of months months)	−0.017*** (0.001)	−0.0004*** (<0.0001)	−0.010*** (0.001)
Photo	0.799*** (0.170)	0.001 (0.006)	0.061 (0.163)
Automatic acceptance	−0.773*** (0.060)	0.131*** (0.002)	3.135*** (0.057)
Hours until departure	−0.039*** (0.0003)	−0.001*** (0.00001)	−0.021*** (0.0003)
Posted since	1.269*** (0.005)	0.011*** (0.0002)	0.292*** (0.004)
Travel time by public transport	1.080*** (0.314)	0.018 (0.011)	−1.519*** (0.299)
Length (# km)	0.007*** (0.001)	−0.0002*** (0.0001)	0.013*** (0.001)
Train strike	4.795*** (0.201)	0.128*** (0.007)	2.949*** (0.191)
Ride description (number of words)	0.033*** (0.002)	0.001*** (0.0001)	0.021*** (0.002)
Constant	13.299*** (0.588)	0.321*** (0.021)	5.671*** (0.560)
Time fixed effects	X	X	X
Route fixed effects	X	X	X
Observations	302,645	317,643	314,361
R^2	0.247	0.075	0.075
Note:			*p<0.1; **p<0.05; ***p<0.01

mation becomes available. This is so because reviews reveal, on average higher quality than expected ex-ante.[33]

To study the impact of reputation, we divide drivers in our dataset into three categories: entrants, defined as drivers with five or fewer reviews, intermediate (with between 6 and 15 reviews), and experienced (with more than 40 reviews). We are interested in measuring the disparity between minority and nonminority drivers in each of these groups. We estimate standard OLS regressions with the same set of controls as in Table 10 for drivers with different levels of experience. The full results are presented in Appendix I; here, we focus on the impact of minority status only.

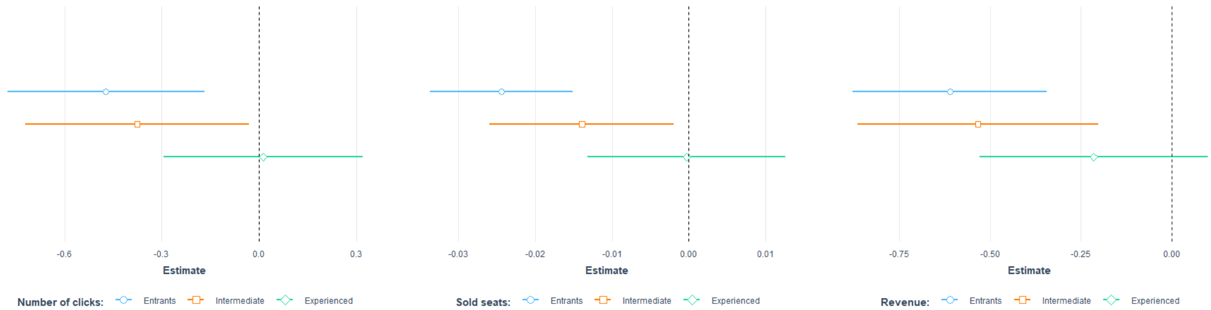
Figure 24 shows the impact of the minority status on the number of clicks (the left panel), the number of sold seats (the center panel), and revenue (the right panel) across various levels of reputation. For entrants (blue), minority status is associated with fewer clicks, sold seats, and lower revenue. The disparity between minority and nonminority drivers decreases with accumulating reviews; it is already smaller at the intermediate level of reputation, and there is no statistically significant difference for drivers with more than 40 reviews.[34]

Controlling for other observables, the initial gap in revenue (for drivers with 0 to 5 reviews) is

[33] In contrast, if discrimination is taste-based, the information about the quality of service provided by minority drivers will not matter. In the taste-based discrimination case, the only relevant information is the ethnicity status itself.

[34] In Appendix K, we show similar patterns using panel data regressions. We also obtain this result using exact and coarsened matching - Appendix J.

Figure 24: Gap between minority and nonminority drivers decreases with reviews



Note: Impact of minority status on: number of clicks (left), sold seats (center), revenue (right) across reputation levels: blue-entrants, orange-intermediate, green - experienced. Coefficients from OLS regressions.

11.8%. It decreases to 6.9% for intermediate drivers (with 6 to 15 reviews) and is as low as 1.6% for experienced drivers (with more than 40 reviews). The results are similar for other measures of performance.[35]

H Output gap: endogeneity of price

In this section, we address the problem of endogeneity of price and quantity in the regression showing minority output gap. Column 1 of the Table 11 introduced the price of the ride in the regression with sold seats as a left-hand side variable (other covariates are unchanged). Column 2 presents an instrumental variables regression, where the price is instrumented with a price of car fuel in the cities of departure and arrival (which we observe on the daily basis), and highway tolls on a given route in a given period.

[35] The gap in the number of clicks is 2.8% for entrant drivers, 2.2% for intermediate drivers, and 0.1% for experienced ones. As to the number of sold seats, the initial gap is 12.2%. It declines to 5.5% with five to fifteen reviews and to 0.1% for experienced drivers.

Table 11: Sold seats: controlling for price and instrumenting it.

	Dependent variable:	
	sold seats	
	OLS	*IV*
	(1)	(2)
minority	−0.013*** (0.003)	−0.006** (0.003)
price	−0.009*** (0.0002)	−0.024*** (0.002)
driver age	−0.001*** (0.0001)	−0.001*** (0.0001)
reviews	0.001*** (0.00005)	0.001*** (0.0001)
reviews2	−0.00000*** (0.00000)	−0.00000*** (0.00000)
male	−0.0002 (0.002)	−0.005** (0.002)
hours untill ride	−0.001*** (0.00001)	−0.001*** (0.00001)
posted since	0.012*** (0.0002)	0.012*** (0.0002)
post per month	−0.005*** (0.001)	−0.005*** (0.001)
length bio	0.0002 (0.0001)	−0.0002* (0.0001)
car price	0.0002 (0.0002)	0.001** (0.0002)
public transport ratio	10.022*** (3.365)	17.286*** (3.499)
km	0.001*** (0.00005)	0.002*** (0.0001)
day	0.018*** (0.004)	0.015*** (0.004)
night	−0.054*** (0.006)	−0.050*** (0.007)
train strike	0.131*** (0.007)	0.171*** (0.006)
length ride (# words)	0.001*** (0.0001)	0.0005*** (0.0001)
picture	0.001 (0.006)	0.002 (0.006)
automatic acceptance	0.114*** (0.002)	0.086*** (0.004)
weekday	−0.042*** (0.004)	−0.041*** (0.004)
day*weekday	0.009* (0.005)	0.012** (0.005)
night*weekday	0.004 (0.008)	0.004 (0.008)
Constant	0.202*** (0.041)	0.046 (0.043)
Observations	318,420	287,754
R^2	0.078	0.064
Note:		*p<0.1; **p<0.05; ***p<0.01

I Reputation effect

Table 12: Revenue regressed over driver and ride characteristics

	Dependent variable:		
	revenue		
	(1)	(2)	(3)
minority	-0.623^{***} (0.142)	-0.451^{**} (0.178)	-0.233 (0.168)
driver age	-0.033^{***} (0.004)	-0.026^{***} (0.005)	-0.008 (0.005)
reviews	0.234^{***} (0.029)	0.078^{***} (0.021)	0.014^{***} (0.001)
male	-0.278^{**} (0.108)	-0.198 (0.131)	0.213 (0.151)
hours till ride	-0.015^{***} (0.0005)	-0.019^{***} (0.001)	-0.028^{***} (0.001)
posted since	0.215^{***} (0.007)	0.304^{***} (0.009)	0.345^{***} (0.009)
post per month	0.043 (0.040)	-0.133^{***} (0.049)	-0.409^{***} (0.034)
seniority (# months)	-0.006^{***} (0.002)	-0.016^{***} (0.003)	-0.029^{***} (0.003)
length bio	0.009 (0.005)	0.010 (0.006)	-0.00002 (0.007)
car price	-0.018^{*} (0.010)	-0.006 (0.012)	-0.019 (0.012)
competition	0.006^{***} (0.002)	0.004^{*} (0.002)	0.005^{***} (0.002)
duration public transport	-0.192 (0.534)	-0.930 (0.678)	-2.719^{***} (0.804)
km	0.003 (0.003)	0.010^{***} (0.004)	0.017^{***} (0.004)
day	0.402^{**} (0.194)	0.770^{***} (0.239)	0.574^{**} (0.246)
night	-1.025^{***} (0.299)	-0.860^{**} (0.391)	-1.668^{***} (0.378)
train strike	2.757^{***} (0.294)	2.981^{***} (0.358)	3.107^{***} (0.547)
length ride (# words)	0.029^{***} (0.004)	0.018^{***} (0.004)	0.013^{***} (0.004)
picture	0.199 (0.278)	0.091 (0.432)	-1.296^{***} (0.413)
automatic acceptance	3.381^{***} (0.107)	3.126^{***} (0.126)	2.929^{***} (0.126)
weekday	-0.509^{**} (0.202)	-0.222 (0.248)	-1.173^{***} (0.244)
travel cost	0.017^{**} (0.008)	0.005 (0.011)	0.020 (0.012)
median revenue	-0.032 (0.130)	-0.061 (0.161)	-0.184 (0.184)
day*weekday	0.096 (0.238)	-0.237 (0.291)	0.243 (0.293)
night*weekday	0.013 (0.368)	-0.612 (0.474)	0.359 (0.454)
Constant	3.745^{**} (1.894)	4.438^{*} (2.466)	15.360^{***} (3.153)
Driver effects	X	X	X
Ride effects	X	X	X
Time effects	X	X	X
Trip effects	X	X	X
Observations	82,563	65,013	68,505
R^2	0.060	0.070	0.096
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 13: Sold seats regressed over driver and ride characteristics

	Dependent variable:		
	taken_seats		
	(1)	(2)	(3)
minority	−0.024*** (0.005)	−0.014** (0.006)	−0.0003 (0.007)
driver age	−0.001*** (0.0001)	−0.001*** (0.0002)	−0.0004** (0.0002)
reviews (#)	0.008*** (0.001)	0.004*** (0.001)	0.001*** (0.00004)
male	−0.005 (0.004)	−0.002 (0.004)	0.011* (0.006)
seniority (# months)	−0.0002** (0.0001)	−0.001*** (0.0001)	−0.001*** (0.0001)
hours till ride	−0.001*** (0.00002)	−0.001*** (0.00002)	−0.001*** (0.00003)
posted since	0.008*** (0.0002)	0.012*** (0.0003)	0.014*** (0.0004)
post per month	0.001 (0.001)	−0.007*** (0.002)	−0.020*** (0.001)
length bio	0.0001 (0.0002)	0.0001 (0.0002)	−0.0001 (0.0003)
car price	−0.0002 (0.0003)	0.0002 (0.0004)	−0.00003 (0.0004)
competition	0.0002*** (0.0001)	0.0001** (0.0001)	0.0002** (0.0001)
public transport ratio	10.958** (5.263)	1.612 (6.970)	10.991 (9.310)
km	−0.00004 (0.0001)	−0.0001 (0.0001)	0.00003 (0.0001)
day	0.013** (0.006)	0.024*** (0.008)	0.019** (0.010)
night	−0.044*** (0.010)	−0.030** (0.013)	−0.079*** (0.015)
train strike	0.103*** (0.010)	0.139*** (0.013)	0.151*** (0.022)
length ride(# words)	0.001*** (0.0001)	0.001*** (0.0001)	0.001*** (0.0001)
picture	−0.002 (0.009)	−0.007 (0.014)	−0.026* (0.015)
automatic acceptance	0.134*** (0.004)	0.125*** (0.004)	0.134*** (0.005)
weekday	−0.029*** (0.007)	−0.028*** (0.009)	−0.074*** (0.009)
day*weekday	0.010 (0.008)	−0.002 (0.010)	0.025** (0.011)
night*weekday	0.005 (0.012)	−0.024 (0.016)	0.022 (0.018)
Constant	0.137** (0.060)	0.279*** (0.084)	0.470*** (0.124)
Observations	91,870	72,597	76,999
R^2	0.066	0.066	0.083
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 14: Number of clicks regressed over driver and ride characteristics

	Dependent variable:		
	number of clicks		
	(1)	(2)	(3)
minority	−0.472*** (0.155)	−0.376** (0.176)	0.012 (0.157)
driver age	−0.074*** (0.004)	−0.062*** (0.005)	−0.033*** (0.005)
reviews (#)	−0.025 (0.031)	0.052** (0.020)	0.017*** (0.001)
male	−1.490*** (0.118)	−1.741*** (0.129)	−0.917*** (0.140)
seniority (# months)	−0.007*** (0.002)	−0.016*** (0.003)	−0.044*** (0.003)
hours till ride	−0.034*** (0.001)	−0.038*** (0.001)	−0.045*** (0.001)
posted since	1.195*** (0.008)	1.348*** (0.010)	1.210*** (0.009)
post per month	−0.186*** (0.044)	−0.470*** (0.048)	−0.780*** (0.032)
length bio	0.002 (0.006)	0.005 (0.006)	−0.004 (0.006)
car price	0.013 (0.010)	0.016 (0.012)	−0.0004 (0.011)
competition	0.011*** (0.002)	0.009*** (0.002)	0.012*** (0.002)
public transport ratio	621.914*** (172.633)	462.974** (201.728)	210.319 (227.733)
km	0.016*** (0.002)	0.016*** (0.002)	0.014*** (0.003)
day	−0.711*** (0.209)	0.494** (0.236)	0.629*** (0.227)
night	−0.078 (0.328)	0.593 (0.386)	−1.286*** (0.351)
train strike	5.235*** (0.330)	4.815*** (0.364)	4.839*** (0.533)
length ride (# words)	0.052*** (0.004)	0.037*** (0.004)	0.011*** (0.003)
picture	1.228*** (0.282)	0.236 (0.389)	−0.883** (0.355)
automatic acceptance	−0.194* (0.116)	−0.454*** (0.124)	−1.561*** (0.118)
weekday	−1.253*** (0.218)	−0.146 (0.245)	−0.894*** (0.225)
day*weekday	1.395*** (0.257)	0.059 (0.287)	0.664** (0.271)
night*weekday	−0.160 (0.403)	−0.591 (0.469)	0.846** (0.423)
Constant	6.853*** (1.990)	7.978*** (2.446)	18.422*** (3.060)
Observations	87,004	69,163	73,834
R^2	0.250	0.259	0.254
Note:			*p<0.1; **p<0.05; ***p<0.01

J Matching Analysis

This project, likewise most in the literature, uses non-experimental data for evaluating the impact of minority status. Hence, estimates of the impact of being a minority may suffer from a bias of the selection on the non-observables. There is a growing, mostly theoretical, literature on the use of matching techniques to address this issue. Rosenbaum and Rubin (1983) and Heckman et al. (1997) demonstrate that this bias can be greatly reduced by use of various matching techniques. Some of their properties are discussed by Abadie and Imbens (2016). A similar methodology has been applied in Sarsons (2017). [36]

The objective of matching exercise is to test the robustness of results from the standard OLS of Section G. We will firstly estimate propensity scores for each of the observations and discard these with extreme values. Secondly, we will perform matching of the minority and non-minority subsamples on driver-specific variables. We will execute both exact matching and coarsened matching. Finally, we will regress model using the matched sample, controlling for listing-specific characteristics.

The propensity score is a logistic regression with minority status being dependent variables and following controls: the price of a car, driver's age, number of posts per month, picture dummy, length of biography, gender, fuel consumption of the car and whether the driver is talkative. The results are displayed in Table 15. Minority drivers are more likely to be a young male and to enjoy conversations. They have on average more expensive cars that consume more fuel; their profiles are also shorter. We delete 5% smallest and 5% largest propensity scores, in this way we delete observations for which we are unlikely to find a counterpart.

J.1 Exact matching

is performed on all driver's characteristics for which we have estimated the logistic regression. In our sample, it means that we have 8809 minority drivers matched with 22617 non-minority drivers. As entrants, we will label minority drivers with less than five reviews and as incumbents (experienced users) these with more than 50 reviews. In the case of exact matching, the definition of an incumbent is extended to drivers with more than 30 reviews so as to increase the size of the group. From Table 16 we can see that even after the matching procedure, minority entrant drivers are facing discrimination. We repeat the same process for drivers with reputation. The results in Table 17 reveal the reputation effect; minority status when users are experienced is insignificant for all measures of economic outcome.

[36] We use matching software developed by Iacus et al. (2009).

Table 15: Propensity score table

	Dependent variable:
	minority
car price	0.024*** (0.001)
driver age	-0.033^{***} (0.001)
post per month	0.058*** (0.003)
picture	0.078*** (0.023)
length bio	-0.013^{***} (0.0005)
male	0.952*** (0.019)
consumption	0.140*** (0.010)
driver blabla	0.329*** (0.014)
Constant	-2.965^{***} (0.063)
Observations	195,333
Log Likelihood	$-75,005.150$
Akaike Inf. Crit.	150,028.300
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 16: Economic outcomes of entrants, exact matching

	Dependent variable		
	number of clicks	Revenue	Taken seats
minority	-1.2546*** (0.341)	-0.69684** (0.221)	-0.0258*** (0.007)
hours until ride	-0.0107 (0.009)	-0.0109 (0.006)	-0.0005* (0.0002)
posted since	2.0561 *** (0.225)	0.2721 (0.145)	0.0059 (0.005)
competition	0.0249 ***(0.005)	0.003*** (0.569)	0.0005*** (0.0001)
day	0.3792 (0.349)	0.5492* (0.226)	0.0121 (0.012)
night	0.8578 (0.517)	-1.2120*** (0.335)	-0.0492 *** (0.005)
notice	-0.4786^{*} (0.122)	-0.0027 (0.145)	0.0031 (0.023)
Matched Observations	19,112		
Note: Trip fixed effects not reported			*p<0.05; **p<0.01; ***p<0.001

Table 17: Economic outcomes of incumbent drivers, exact matching

	Dependent variable		
	number of clicks	Revenue	Taken seats
minority	-0.2876 (0.383)	0.1409 (0.303)	0.0106 (0.012)
hours untill ride	-0.0340** (0.0117)	0.0112 (0.009)	0.0009* (0.0004)
posted since	1.8896*** (0.282)	1.341*** (0.223)	0.0626 *** (0.009)
competition	0.0168 **(0.005)	0.0076 (0.004)	0.0003 (0.0002)
day	0.1916 (0.394)	0.0322 (0.311)	0.0177 (0.012)
night	0.5573 (0.634)	-1.9742*** (0.500)	-0.0829*** (0.019)
notice	-0.1398 (0.280)	-0.8167 *** (0.221)	-0.0436 *** (0.008)
Matched Observations	12314		

Note: Trip fixed effects not reported *p<0.05; **p<0.01; ***p<0.001

J.2 Coarsened Matching

Coarsened Matching is a method used to increase the number of matched observations. We introduce bins in which we will match non-binary covariates: age of the driver, the price of a car, number of posts per month, length of bio and fuel consumption of the car. Choice of cutoffs influences the precision of matching procedure as well as the number of matched observations; we match within a quartile for each of the variables. In this way, we match 14146 minority drivers with 45959 nonminority ones, which is almost a twofold increase. We present only the coefficient of minority status (Table 18). In this coarsely matched sample, we also see a clear reputation effect. Minority entrants have lower economic outcomes, however after they build reputation most of the effect goes away. These results depend on cut-offs for labeling as entrants/ incumbents, as well as on the selection of bins for coarsened matching; they are however robust to local changes.

Table 18: Economic outcomes entrants and incumbents, coarsened matching

	Dependent variable		
	number of clicks	Revenue	Taken seats
minority (entrant)	-0.9276 *** (0.271)	-0.6032*** (0.180)	-0.0191** (0.006)
minority (incumbent)	-0.3883 (0.303)	0.0330 (0.229)	-0.0084 (0.010)
Matched Observations (both models)	57,853		

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

Table 19: Propensity score

	Dependent variable:
	f minority
car price	0.024^{***} (0.001)
driver age	-0.033^{***} (0.001)
post per month	0.058^{***} (0.003)
picture	0.078^{***} (0.023)
length bio	-0.013^{***} (0.0005)
male	0.952^{***} (0.019)
consumption	0.140^{***} (0.010)
driver blabla	0.329^{***} (0.014)
Constant	-2.965^{***} (0.063)
Observations	195,333
Log Likelihood	−75,005.150
Akaike Inf. Crit.	150,028.300
Note:	*p<0.1; **p<0.05; ***p<0.01

K Panel data results

Thousands of drivers are active on BlaBlaCar at any moment; thus, every time we collect data, we observe only a fraction of all available listings. As a consequence, we see most drivers only once. However, in some cases (22.800 drivers), we see the driver at least twice, which gives us a panel with almost 56.800 observations. However, this sample is unbalanced, with drivers being observed between 2 and 30 times. We use several standard models that allow us to compare the gap associated with being a minority entrant or incumbent entrant. Reduction in the sample size results in lower significance of our estimates. However, the signs and point estimates appear to confirm our hypothesis.

We estimate the following model:

$$y_{it} = \alpha + X_{it}\beta + Z_i\gamma + c_i + \tau_t + \epsilon_{it}$$

where c_i are individual fixed effects and ϵ_{it} is an idiosyncratic error term.

We present minority dummies and the products of minority and entrant dummies. Similarly to the cross-sectional analysis in the main body of the paper, we conclude that upon entering the market, minority drivers receive lower outcomes and that this effect weakens as drivers receive reviews. Again, the reputation effect is significant for all measures of economic performance.

	Dependent variable: number of clicks		
	Pooled	Between	Random
minority	0.288 (0.202)	0.409 (0.275)	0.317 (0.236)
entrant	−0.995*** (0.143)	−0.811*** (0.179)	−0.764*** (0.155)
minority*entrant	−0.678* (0.353)	−0.692 (0.449)	−0.717* (0.387)
driver's age	−0.036*** (0.005)	−0.038*** (0.006)	−0.036*** (0.006)
talkative	0.220* (0.123)	0.363** (0.156)	0.282** (0.141)
male	−1.074*** (0.142)	−1.105*** (0.171)	−1.128*** (0.159)
hours until ride	−0.028*** (0.0005)	−0.023*** (0.001)	−0.029*** (0.0005)
posted since	1.136*** (0.010)	1.068*** (0.016)	1.172*** (0.010)
bio (# words)	−0.002 (0.004)	−0.003 (0.005)	−0.002 (0.004)
car price	−0.018 (0.012)	−0.031** (0.015)	−0.021 (0.014)
competition	0.036*** (0.002)	0.035*** (0.003)	0.034*** (0.002)
median revenue	−0.00002 (0.00003)	−0.0001* (0.00004)	−0.00000 (0.00003)
public transport ratio	−0.909 (7.222)	−1.761 (10.861)	−2.131 (7.765)
km	0.007*** (0.0004)	0.006*** (0.001)	0.006*** (0.0004)
day	0.538** (0.231)	0.574 (0.364)	0.462** (0.231)
night	−0.605* (0.357)	−1.134* (0.581)	−0.763** (0.358)
train strike	3.269*** (0.325)	3.049*** (0.538)	3.545*** (0.319)
ride (# words)	0.018*** (0.002)	0.021*** (0.002)	0.020*** (0.002)
picture	0.246 (0.201)	0.494* (0.260)	0.496** (0.230)
automatic acceptance	−1.334*** (0.122)	−1.299*** (0.164)	−1.307*** (0.132)
weekday	−0.018 (0.236)	−0.457 (0.387)	0.112 (0.237)
consumption	0.278*** (0.084)	0.377*** (0.106)	0.303*** (0.095)
day*weekday	0.465 (0.284)	0.925** (0.460)	0.389 (0.284)
night*weekday	−0.018 (0.444)	1.577** (0.746)	0.094 (0.443)
Constant	11.748*** (0.880)	10.465*** (1.234)	11.477*** (0.948)
Observations	56,760	22,794	56,760
R^2	0.244	0.220	0.262
Note:			*p<0.1; **p<0.05; ***p<0.01

	Dependent variable: sold seats		
	Pooled	Between	Random
minority	0.002 (0.009)	0.016 (0.011)	0.002 (0.009)
entrant	−0.060*** (0.011)	−0.058*** (0.012)	−0.059*** (0.011)
minority*entrant	−0.035** (0.016)	−0.041** (0.018)	−0.035** (0.016)
male	0.005 (0.008)	0.004 (0.009)	0.004 (0.008)
driver's age	−0.0004* (0.0002)	−0.0003 (0.0002)	−0.0004* (0.0002)
talkative	0.001 (0.005)	0.003 (0.006)	0.001 (0.006)
hours until ride	−0.001*** (0.00002)	−0.001*** (0.00003)	−0.001*** (0.00002)
posted since	0.016*** (0.0004)	0.012*** (0.001)	0.016*** (0.0004)
bio (# words)	−0.0001 (0.0002)	−0.0003 (0.0002)	−0.0001 (0.0002)
car price	−0.0003 (0.001)	−0.001 (0.001)	−0.0004 (0.001)
competition	0.001*** (0.0001)	0.001*** (0.0001)	0.001*** (0.0001)
median revenue	0.00000*** (0.00000)	0.00001*** (0.00000)	0.00000*** (0.00000)
public transport ratio	−0.146 (0.318)	−0.566 (0.440)	−0.147 (0.322)
km	−0.00002 (0.00002)	−0.0001** (0.00002)	−0.00002 (0.00002)
day	0.015 (0.010)	0.004 (0.015)	0.015 (0.010)
night	−0.048*** (0.016)	−0.062*** (0.023)	−0.048*** (0.016)
train strike	0.126*** (0.014)	0.110*** (0.022)	0.128*** (0.014)
ride (# words)	0.0004*** (0.0001)	0.0004*** (0.0001)	0.0004*** (0.0001)
picture	0.002 (0.009)	0.015 (0.010)	0.003 (0.009)
automatic acceptance	0.109*** (0.005)	0.108*** (0.007)	0.109*** (0.005)
weekday	−0.045*** (0.010)	−0.059*** (0.016)	−0.045*** (0.010)
consumption	0.020*** (0.004)	0.021*** (0.004)	0.020*** (0.004)
minority*entrant	−0.035** (0.016)	−0.041** (0.018)	−0.035** (0.016)
entrant*male	−0.016 (0.013)	−0.022 (0.014)	−0.015 (0.013)
day*weekday	0.019 (0.012)	0.034* (0.019)	0.019 (0.013)
night*weekday	−0.020 (0.019)	0.012 (0.030)	−0.020 (0.020)
Constant	0.180*** (0.039)	0.158*** (0.050)	0.175*** (0.039)
Observations	59,359	23,076	59,359
R^2	0.089	0.085	0.088
Note:			*p<0.1; **p<0.05; ***p<0.01

	Dependent variable: revenue		
	Pooled	Between	Random
minority	−0.334 (0.213)	0.022 (0.275)	−0.272 (0.228)
entrant	−1.387*** (0.150)	−1.452*** (0.179)	−1.308*** (0.155)
minority*entrant	−0.680* (0.372)	−0.866* (0.448)	−0.741* (0.387)
driver's age	−0.006 (0.005)	−0.002 (0.006)	−0.005 (0.005)
talkative	0.020 (0.129)	0.065 (0.155)	0.026 (0.137)
male	−0.201 (0.148)	−0.307* (0.170)	−0.240 (0.156)
hours untill ride	−0.018*** (0.0005)	−0.016*** (0.001)	−0.019*** (0.0005)
posted since	0.371*** (0.010)	0.290*** (0.014)	0.375*** (0.010)
bio (# words)	−0.001 (0.004)	−0.005 (0.005)	−0.001 (0.004)
car price	−0.007 (0.013)	−0.022 (0.015)	−0.010 (0.013)
competition	0.024*** (0.002)	0.024*** (0.003)	0.024*** (0.002)
median revenue	0.0002*** (0.00003)	0.0003*** (0.00004)	0.0002*** (0.00003)
public transport ratio	−33.375*** (7.569)	−40.181*** (10.934)	−33.318*** (7.835)
km	0.013*** (0.0004)	0.011*** (0.001)	0.013*** (0.0004)
day	0.410* (0.243)	0.445 (0.367)	0.400 (0.244)
night	−1.341*** (0.373)	−2.091*** (0.579)	−1.300*** (0.376)
train strike	2.367*** (0.339)	1.779*** (0.543)	2.429*** (0.338)
ride (# words)	0.007*** (0.002)	0.008*** (0.002)	0.007*** (0.002)
picture	0.087 (0.212)	0.366 (0.260)	0.170 (0.225)
automatic acceptance	2.064*** (0.128)	2.012*** (0.164)	2.104*** (0.133)
weekday	−0.847*** (0.249)	−1.142*** (0.390)	−0.828*** (0.251)
consumption	0.315*** (0.088)	0.340*** (0.106)	0.325*** (0.093)
minority*entrant	−0.680* (0.372)	−0.866* (0.448)	−0.741* (0.387)
day*weekday	0.317 (0.299)	0.446 (0.465)	0.290 (0.300)
night*weekday	−0.215 (0.465)	0.938 (0.745)	−0.232 (0.467)
Constant	−1.089 (0.926)	−2.287* (1.239)	−1.200 (0.957)
Observations	58,621	23,018	58,621
R^2	0.095	0.093	0.094
Note:			*p<0.1; **p<0.05; ***p<0.01

L Strikes

Table 20 presents means of selected characteristics of drivers on days of strike and days without a strike. Subset of drivers active in the period 03/04/2018 to 28/06/2018.

Results of the main specification with number of sold seats as the dependent variable Table21.

Alternative definition of treated: minority drivers with less than three reviews driving on the day of strike (table 22). We see a higher significance on the treated status.

Table 20: Characteristics of drivers on days of strike and non-strike days.

	Minority	Male	Reviews	Notice	Age	Car value	Published rides	Posts per month	Reputation	Km
No strike	0.1483	0.7258	28.5598	21.5242	37.5912	6.1926	31.2066	1.6434	4.6398	432.0588
Strike	0.1469	0.7291	28.4609	22.3024	38.0868	6.1491	31.4207	1.6502	4.6397	426.7940

Note: Means of selected variables

Table 21: Difference in differences estimation sold seats as a dependent variable

	Dependent variable:		
	sold seats		
	(1)	(2)	(3)
treated	-0.042^{***} (0.013)	-0.024 (0.015)	-0.022 (0.015)
after	-0.154 (0.135)	-0.151 (0.148)	-0.163 (0.148)
did	0.062^{***} (0.023)	0.050^{*} (0.026)	0.050^{*} (0.026)
minority	-0.012^{***} (0.003)	-0.006^{*} (0.004)	-0.003 (0.004)
Driver characteristics			x
Listing characteristics		x	x
Route effects	x	x	x
Time effects	x	x	x
Observations	300,636	243,407	243,407
R^2	0.032	0.033	0.035
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 22: Difference in differences estimation with revenue as dependent variable

	Dependent variable: revenue		
	(1)	(2)	(3)
treated	-0.765 (0.539)	-0.831 (0.726)	-0.925 (0.725)
after	-3.594 (3.671)	-3.805 (4.020)	-3.998 (4.018)
did	0.952^{*} (0.577)	1.304^{**} (0.620)	1.340^{**} (0.620)
minority	-0.556^{***} (0.085)	-0.382^{***} (0.097)	-0.316^{***} (0.097)
Driver characteristics			x
Listing characteristics		x	x
Route effects	x	x	x
Time effects	x	x	x
Observations	297,006	240,473	240,473
R^2	0.040	0.042	0.043
			*p<0.1; **p<0.05; ***p<0.01

Note:Treated: minority drivers with less than 3 reviews

M Proofs

M.1 Proof of Proposition 1:

The optimal level of effort satisfies the first order condition of driver's maximization problem 1, which writes $\sum_{s=t}^{\infty}\delta^{s-t}\mathbf{E}_t\left[\frac{\partial\pi_s}{\partial a_t}\right]-g'(a_t)=0$.

First, we calculate the derivative of per-period profits at $s>t$ with respect to effort at t (note that the effort of period t does not influence profits in earlier periods):

$$\frac{\partial\pi_s}{\partial a_t}=M_s\left(\frac{\partial\mathcal{S}_s}{\partial\hat{q}_s}\frac{\partial\hat{q}_s}{\partial a_t}+\frac{\partial\mathcal{S}_s}{\partial p_s}\frac{\partial p_s}{\partial a_t}\right)(p_s-c)+M_s\frac{\partial p_s}{\partial a_t}\mathcal{S}_t=M_s\frac{\partial\mathcal{S}_s}{\partial\hat{q}_s}\frac{\partial\hat{q}_s}{\partial a_t}(p_s-c)\tag{18}$$

,where $\hat{q}_d=\mathbf{E}_t[q_d|h_d]$. The second equality is due to driver's static price optimization.

Given the assumption on the driver's choice problem and the utility function we can represent the probability of a passenger choosing a seat of driver i as

$$s_{it}=\frac{\exp(\alpha\mathbf{E}_t[q_{it}|h_{it}]+\beta\mathbf{Var}[q_{it}|h_{it}]+\gamma p_{it})}{1+\sum_{k=1}^{N}(\exp(\alpha\mathbf{E}_t[q_{kt}|h_{kt}]+\beta\mathbf{Var}[q_{kt}|h_{kt}]+\gamma p_{kt})}.\tag{19}$$

Therefore, the elasticity of demand with respect to expected quality writes $\frac{\partial\mathcal{S}_{it}}{\partial\hat{q}_{it}}=M_t\alpha\mathcal{S}_{it}(1-\mathcal{S}_{it})$.

Observe that the marginal effect on effort on perceived output for $s>t$ simplifies to:

$$\frac{\partial\hat{q}_{is}}{\partial a_{it}}=\frac{\partial}{\partial a_t}\left\{\frac{\tau_g\hat{\mu}_g}{\tau_{gs}}+\frac{\tau_\epsilon}{\tau_{mk}}\sum_{d=1}^{s-1}(\eta_i+a_{id}-\mathbf{E}_d\left[a^*_{id}(h_{id})\right])+\mathbf{E}_s\left[a^*_d(h_{is})\right]\right\}=\frac{\tau_\epsilon}{\tau_{gs}}\tag{20}$$

,where $\tau_{gs}=\tau_g+(s-1)\tau_\epsilon$ and $a^*_{it}(h_{it})$ is the anticipated, by passengers, level of effort. We can therefore simplify equation 18, so that:

$$\frac{\partial\pi_{is}}{\partial a_{it}}=M_s\alpha\mathcal{S}_{id}(1-\mathcal{S}_{is})\frac{\tau_\epsilon}{\tau_{gs}}(p_{is}-c_i),\tag{21}$$

consequently the equilibrium level of effort is period t is given by

$$a_{it} = (g')^{-1}\left(\sum_{s=t+1}^{\infty}\delta^{s-t}M_d\alpha\mathcal{S}_{id}(1-\mathcal{S}_{is})\frac{\tau_\epsilon}{\tau_{gs}}(p^*_{is}-c_i)\right). \tag{22}$$

Both δ^{s-t} and $\frac{\tau_\epsilon}{\tau_{gs}}$ tend asymptotically to zero. The equilibrium level of prices p^*_{id} is shown by proposition 2 to tend to static prices under full information, finally as the consequence of proposition 2, $\mathcal{S}_{id}$ also tends to static market share under full information. For the formal proof see proof of proposition 2, the intuition is that as drivers receive the evaluations their type becomes fully revealed, therefore, they have no further incentive to strategically deviate from short term profit maximizing prices. Thus, all terms are bounded and the right hand side of 23 converges to 0 as t goes to infinity.

M.2 Proof of Corollary 1:

First part: Note, that we define the discrimination keeping the level of prices fixed. Thus, it's enough to show how changes in the level of discrimination impact the effort level through market shares. The optimal level of effort writes

$$a_{it} = g^{-1'}\left(\sum_{d=t+1}^{\infty}\delta^{d-t}M_d\alpha s_{id}(1-s_{id})\frac{\tau_\epsilon}{\tau_{gd}}(p^*_{id}-c_i)\right). \tag{23}$$

, while the discrimination is defined as: $D(h,p) \equiv \mathbf{E}\left[\mathcal{S}(p,h)|n\right] - \mathbf{E}\left[\mathcal{S}(p,h)|m\right]$. An increase in $D(h,p)$ comes either from an increase in the number of sold seats of a nonminority driver or a decrease by a minority driver. As both drivers are in the same market, the two possible changes result in a decrease in $s_{id}(1-s_{id})$, where i is the minority driver, as long as $s_{id} \in [0,1/2]$, which is the case we consider.

Second part: Again we consider the discrimination as the difference in the number of sold seats keeping the level of prices and review histories fixed. Recall that the number of sold seats writes

$$\mathcal{S}_{it}(p,h) = M_t\frac{\exp(\alpha\mathbf{E}[q_{it}|h_{it}] + \beta\mathbf{Var}[q_{it}|h_{it}] + \gamma p_{it})}{1+\sum_{k=1}^{N}(\exp(\alpha\mathbf{E}[q_{kt}|h_{kt}] + \beta\mathbf{Var}[q_{kt}|h_{kt}] + \gamma p_{kt})} \tag{24}$$

and that the expected quality writes $\mathbf{E}[q_{it}|h_{it}] = \frac{\tau_g\hat{\mu}_g}{\tau_g+t\tau_\epsilon}$. As driver receives reviews the expected effort goes to zero, so the expected quality converges to driver's type,($\lim_{t\to\infty}\mathbf{E}[q_{it}|h_{it}] = \eta_i$). For the same level of discrimination at period t, drivers expected type is higher when discrimination is associated with a negatively biased belief-based partiality than an unbiased belief based partiality. Consequently, the discounted sum of profit is higher under negatively biased belief-based partiality, and so are the efforts by equation 23.

M.3 Proof of Proposition 2

As driver receives reviews the beliefs about the expected quality and about the variance of it are revise,

$$\lim_{t\to\infty}\{\mathbf{E}[q_{it}|h_{it}]\} = \lim_{t\to\infty}\left\{\frac{\tau_g+\hat{\mu}_g}{\tau_g+t\tau_\epsilon}+\frac{\tau_\epsilon}{\tau_g+t\tau_\epsilon}\sum_{s=1}^{t}(\eta_i+\epsilon_s)\right\} = \eta_i \lim_{t\to\infty}\left\{\frac{t}{\frac{\tau_g}{\tau_\epsilon}+t}\right\} = \eta_i. \tag{25}$$

By analogous argument, the variance of expected quality shrinks to the variance of quality reports,

$$\lim_{t\to\infty}\{\mathbf{Var}\left[\mathbf{E}[q_{it}|h_{it}]\right]\} = \lim_{t\to\infty}\left\{\mathbf{E}\left[\mathbf{E}[q_{it}]-\mathbf{E}[q_{it}|h_{it}]^2\right]|q_{it}\right\} = \frac{1}{\tau_\epsilon}. \tag{26}$$

M.4 Proof of Proposition 3

Recall that a discrimination is defined as $D(h,p)\equiv \mathbf{E}\left[\mathcal{S}(p,h)|n\right]-\mathbf{E}\left[\mathcal{S}(p,h)|m\right]$, we want to show that $D(h,p)$ between a driver i from n and j from m goes to zero when $\eta_i=\eta_j$.

$$\begin{aligned}
&\lim_{t\to\infty} D(h,p) = \\
&\lim_{t\to\infty}\left\{M_t\frac{\exp\left(\alpha\mathbf{E}_t[q_{it}|h_{it}]+\beta\mathbf{Var}[q_{it}|h_{it}]+\gamma p^*_{it}\right)}{1+\sum_{k=1}^{N}\exp(\alpha\mathbf{E}_t[q_{kt}|h_{kt}]+\beta\mathbf{Var}[q_{kt}|h_{kt}]+\gamma p^*_{kt})} - M_t\frac{\exp\left(\alpha\mathbf{E}_t[q_{jt}|h_{jt}]+\beta\mathbf{Var}[q_{jt}|h_{jt}]+\gamma p^*_{jt}\right)}{1+\sum_{k=1}^{N}\exp(\alpha\mathbf{E}_t[q_{kt}|h_{kt}]+\beta\mathbf{Var}[q_{kt}|h_{kt}]+\gamma p^*_{kt})}\right\} \\
&\propto \lim_{t\to\infty}\left\{\exp(\alpha\mathbf{E}_t[q_{it}|h_{it}]+\beta\mathbf{Var}[q_{it}|h_{it}]+\gamma p^*_{it}) - (\exp\alpha\mathbf{E}_t[q_{jt}|h_{jt}]+\beta\mathbf{Var}[q_{jt}|h_{jt}]+\gamma p^*_{jt})\right\}
\end{aligned} \tag{27}$$

,where the proportionality is due to the assumption that $\mathbf{E}[\Omega_t]=\mathbf{\Omega}$, that is the entry process that results in the emergence of a market structure does not depend on the characteristics and actions of individual drivers. Hence, we have that,

$$\begin{aligned}
&\lim_{t\to\infty}\left\{\exp(\alpha\mathbf{E}_t[q_{it}|h_{it}]+\beta\mathbf{Var}[q_{it}|h_{it}]+\gamma p^*_{it}) - (\exp\alpha\mathbf{E}_t[q_{jt}|h_{jt}]+\beta\mathbf{Var}[q_{jt}|h_{jt}]+\gamma p^*_{jt})\right\} = \\
&\exp(\alpha\eta_i+\beta\frac{1}{t_\epsilon}+\gamma\tilde{p}_i)-\exp(\alpha\eta_j+\beta\frac{1}{t_\epsilon}+\gamma\tilde{p}_j)=0.
\end{aligned} \tag{28}$$

The first equality stems from the Proposition 1 and the Proposition 2 and the second one from the fact that $\eta_i=\eta_j$.

N Grades do not depend on prices

We investigate whether grades depend on the prices. We regress price, reputation, plus a full set of other controls on grades obtained. We find that in the OLS estimation there is a positive impact of prices on grades. However, after instrumenting the prices with cost shocks and controlling for driver-specific unobservable effect, we find that the effect is statistically insignificant.

Table 23: Impact of prices on grades

	Dependent variable:	
	grade	
	OLS	*panel IV*
	(1)	(3)
price	0.003^{*} (0.002)	-0.016 (0.067)
reputation	0.655^{***} (0.021)	0.483^{***} (0.076)
Observations	10,828	1,072
Driver FE		x
Driver characteristics	x	x
Time effects	x	x
Route effects	x	x
Listing effects	x	x
Note:	*p<0.1; **p<0.05; ***p<0.01	

Note: regression (1) OLS pooling estimatiot. Regression (2) within driver variation in prices. Prices instrumented with cost shocks: time and space variation in prices and highway tolls

O Random coefficients demand estimation

We assumed that the utility of passengers is fully captured by drivers' observed characteristics and a random component. We can thus form *driver categories* that are demand relevant and can be useful for our inquiry: we divide drivers into categories based on the number of reviews: 0, 1-2, 3-4,5-9, and more than 10, together with a minority status (so a category is, for example, zero reviews and not a minority). We aggregate market shares into these categories: thus assuming that passengers are indifferent between any driver in a category. We use these categories as product IDs in a classical BLP setting; this approach has a valuable feature of mitigating the problem of zero market shares. However, we still have some markets where not all product categories are present. We also introduce a random component on price coefficient. Thus, our demand specification takes the following form:

$$Q_{i,t} = M_t \int \frac{\exp(\alpha \bar{w}_{i,t} + \xi_i + \gamma_{j,t} \bar{p_{i,t}} + \phi_t)}{1 + \sum_k \exp(\alpha \bar{w}_{k,t} + \xi_k + \gamma \bar{p_{k,t}} + \phi_t)} dH(\gamma_{j,t})$$

,where $\bar{w}_{i,t}$ is the average price within a category of drivers, ξ_i is a driver category dummy, and H is the join distribution of passenger heterogeneity in $\gamma_{j,t}$.

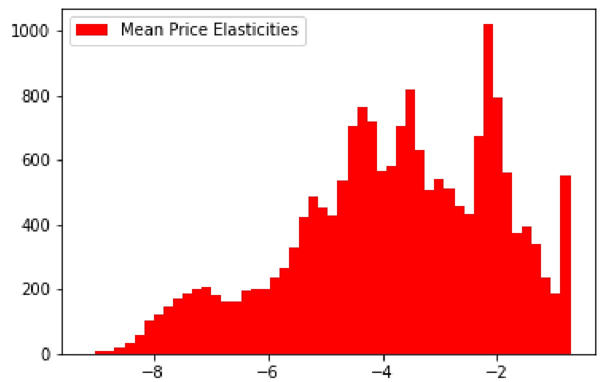
To address the standard problem of the endogeneity of price, we employ two instrumental strategies. First, we use cost-shifters: over time, the price of gas changes, and we can observe the average price at gas stations in any given city on any given day.[37] These prices change over time (because of oil price fluctuations) and location (e.g., due to varying intensities of competition between filling stations). Additionally, the level of highway tolls varies across routes. Second, we observe the characteristics of all drivers available in a given market: we derive measures of isolation in characteristics spaces.

There are many small markets in our dataset; we have more than 64000 markets, with sometimes fewer than five drivers per market. Therefore, we often observe zero market shares. As noted by Gandhi et al. (2013), a typical "fix" in such a case is to add a small ϵ to all market shares or drop observations with zero market share, which effectively lumps them with the outside option. Unfortunately, both methods lead to biased estimates. In the baseline model, we add ϵ to the market shares of all categories. Furthermore, in some categories are missing from some markets, which can be correlated with a trip fixed effects: for example more minority drivers on specific routes.

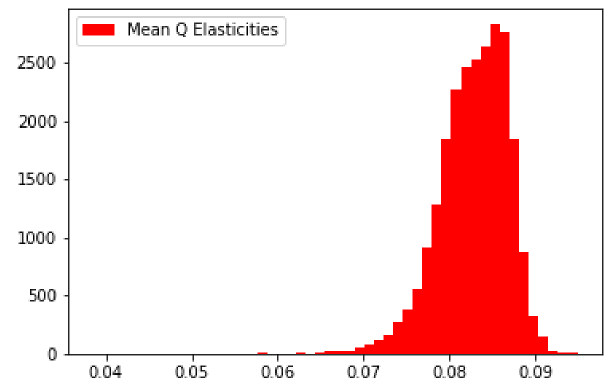
We use Python implementation by Conlon and Gortmaker (2019). Figure 25 shows estimated elasticities with respect to price and quality measure.

Figures 25C and 25D use Reynaert and Verboven (2014) to reweigh instruments. Introducing random coefficient on price does not have a big impact on the magnitude of price elasticity. However, we see that elasticity of demand with respect to price is significantly reduced following the optimal instruments procedure. We conclude that elasticity of price is much higher than that of the quality and that the baseline (standard logit) estimates give a reasonable approximation of the more complex model.

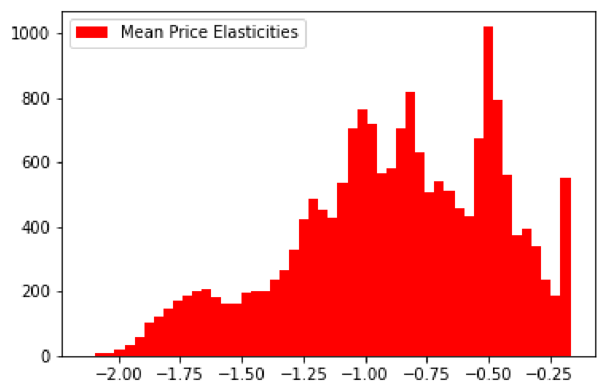
[37] www.prix-carburants.gouv.fr



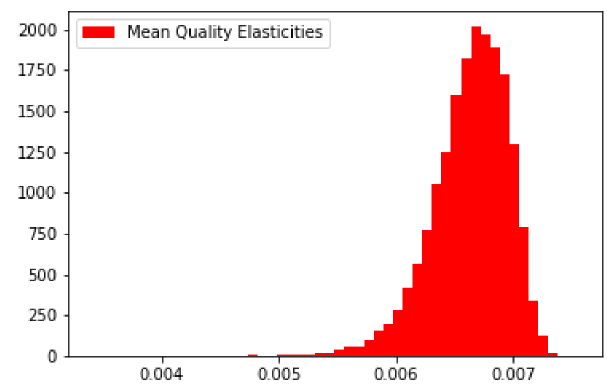
(A) Price elasticity



(B) Quality elasticity



(C) Price elasticity with optimal instruments



(D) Quality elasticity with optimal instruments

Figure 25: Random coefficients logit demand.

P Demand estimation results all variables

	Model 1	Model 2	Model 3
Ride price	-0.00 (0.00)***	-0.00 (0.00)***	-0.00 (0.00)***
Type	0.12 (0.06)*	0.13 (0.06)*	
Log(number reviews)	0.15 (0.02)***	0.14 (0.02)***	0.15 (0.02)***
Automatic acceptance	0.38 (0.04)***	0.38 (0.04)***	0.39 (0.04)***
Picture	0.56 (0.22)*	0.63 (0.23)**	0.63 (0.23)**
Max 2 passengers	-0.19 (0.04)***	-0.20 (0.04)***	-0.21 (0.04)***
Rush time	0.21 (0.06)***	0.24 (0.06)***	0.24 (0.06)***
Day (no rush)	0.08 (0.07)	0.13 (0.07)	0.12 (0.07)
Posted since	0.06 (0.00)***	0.06 (0.00)***	0.06 (0.00)***
Notice	-0.05 (0.00)***	-0.05 (0.00)***	-0.05 (0.00)***
Seniority months	-0.00 (0.00)	-0.00 (0.00)	-0.00 (0.00)
Weekdn	-0.13 (0.04)**	-0.12 (0.04)**	-0.12 (0.04)**
Car price		0.00 (0.00)	0.00 (0.00)
Minority		0.06 (0.05)	
Driver Age		-0.00 (0.00)	
Reputation			0.24 (0.10)*
Time effects	x	x	x
Route effects	x	x	x
AIC	31929.66	30150.03	30145.51
R^2	0.45	0.45	0.45
Max. R^2	0.49	0.49	0.49
Num. events	154259	147905	147905
Num. obs.	470165	442839	442839
Missings	0	0	0

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 24: Demand estimates: subset of markets.

Q Belief updating with discrete reports

The market forms a prior based on driver's characteristics which are observed on her profile, later on as market receives signals about the performance of the driver, beliefs are updated. Holmström (1999) assumes the prior to be normally distributed with mean and variance: $\eta \sim N(m_1, h_1)$; also, he assumes that signals are distributed normally and continously. This leads to a formation of posterior beliefs:

$$\mathbf{E}\left[\eta|z_s\right] = \frac{h_1 m_1 + h_\epsilon \sum_{s_1}^{t} z_s}{h_1 + t h_\epsilon} \tag{29}$$

However, we cannot apply this formula directly because the evaluations are not continuous. Suppose that realizations of output are continuous, but the signals received by the market are discrete. However, there is an objective rule, such that if a realizations falls within a given interval there is always the same grade given: for example, a grade 3 is given when the observed realized output falls within the interval 2.5-3.5, a grade of 5 is given when the observed output is above 4.5. This allows us to calculate marginal probabilities, and characterize the posterior belief, so:

$$\pi(\theta|y) = \frac{f_{y|\theta}(y|\theta)\pi(\theta)}{\int_\Theta f_{y|\theta}(y|\theta)\pi(\theta)d(\theta)} \equiv f_{y|\theta}(y|\theta)\pi(\theta) \tag{30}$$

,where $\pi(\theta|y)$ denotes a probability of being of type θ while getting a grade y and $f_{y|\theta}(y|\theta)$ is a conditional probability of a conditional distribution, the empirical counteraprt of equation (4) is

$$\mathbf{E}\left[\theta|Y=y\right] = \frac{P(Y=y \text{ and } \eta=\theta)}{P(Y=y)} * m_i$$

We are currently improving our estimates to account for this.

R Estimation of the cost of effort function

We are interested in estimating function $g(a_{i,t})$ that measures the cost of exerting effort. The optimal levels of effort, in our model, are determined by a following relation:

$$a_{imt} = \gamma\left(\sum_{s=t}^{n} \beta^{s-t} \frac{h_\epsilon}{h_{mk}} \frac{\alpha}{\gamma} \mathbb{E}[M_k s_{ik}]\right) + \varepsilon_{ijt} \tag{31}$$

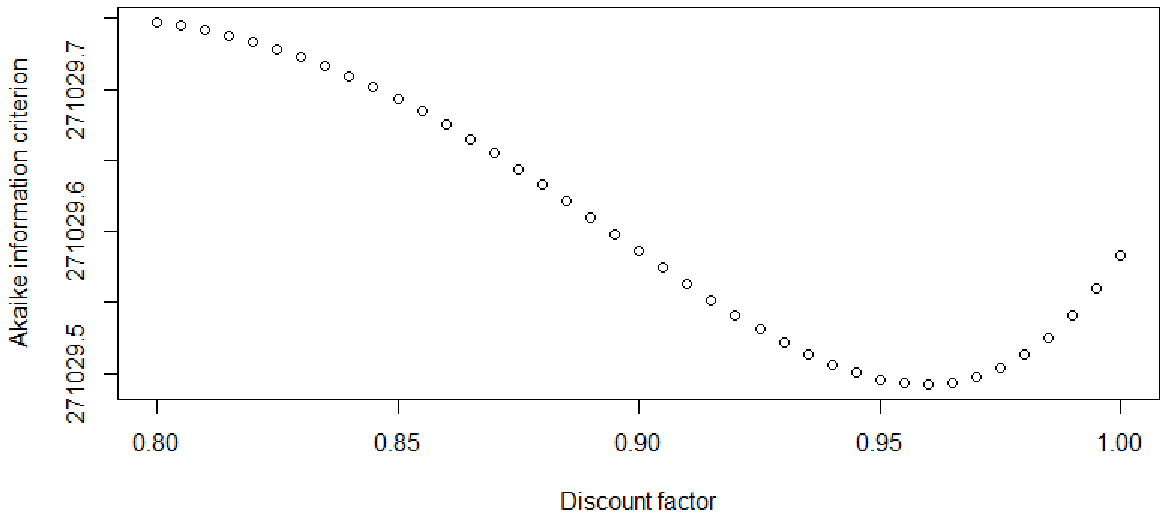


Figure 26: Comparison of the quadratic regression of the cost of effort using different discount factors

,where $\gamma(\cdot) = g^{-1,'}(\cdot)$ in the baseline case cost of effort follows a quadratic function. The discounted sum of profits depends on the discount factor. We estimate the quadratic model for different levels of the discount factor and compare the fit using AIC. Figure 26 shows the AIC on different levels of the discount factor.

The lowest AIC is achieved for the discount factor of 0.96, and the rest of the models are estimated using this discount factor. In the next steps, we fit higher order Chebyshev polynomials on the discounted sum of profits to compare the fit with the quadratic function. Table 25 presents estimates of coefficients of these polynomials

In the next step we present ANOVA results in Table 26 From the ANOVA analysis we conclude that inclusion of 2nd and 3rd degree terms improves fit of the model. Finally we present predictions with confidence intervals for linear, quadratic and 3rd model, see Figure 27. The difference between the quadratic model the 3rd degree polynomial is clear at the high levels of the horizontal axis.

	Dependent variable:				
	effort				
	(1)	(2)	(3)	(4)	(5)
linear	0.97***	−0.62	−12.30***	−14.45	156.12
	(0.05)	(0.47)	(3.17)	(18.97)	(125.35)
2nd degree		4.36***	−12.30***	68.38***	−1,762.95
		(1.29)	(17.23)	(152.66)	(1,351.61)
3rd degree			−113.68***	−174.95	9,656.81
			(30.51)	(535.11)	(18,653.10)
4th degree				79.15	−25,581.11
				(690.23)	(18,653.10)
5th degree					26,304.71
					(19,108.47)
Constant	−0.16***	−0.02	0.67***	0.77	−5.41
	(0.01)	(0.04)	(0.19)	(0.87)	(4.57)
Observations	138,390	138,390	138,390	138,390	138,390
R^2	0.003	0.003	0.003	0.003	0.003
Adjusted R^2	0.003	0.003	0.003	0.003	0.003
Residual Std. Error	0.64 (df = 138388)	0.64 (df = 138387)	0.64 (df = 138386)	0.64 (df = 138385)	0.64 (df = 138384)
F Statistic	426.52*** (df = 1; 138388)	218.98*** (df = 2; 138387)	150.63*** (df = 3; 138386)	112.98*** (df = 4; 138385)	90.76*** (df = 5; 138384)
Note:					*p<0.1; **p<0.05; ***p<0.01

Table 25: Estimates of $g(\cdot)$: Column 1 linear model, columns 2-5 polynomials of increasing degrees

Degree	Res.Df	Sum of Sq	F	Pr(>F)	
1	138388				
2	138387	4.7344	11.4098	0.0007308	***
3	138386	5.7617	13.8858	0.0001943	***
4	138385	0.0055	0.0132	0.9086991	
5	138384	0.7863	1.8950	0.1686383	

Table 26: ANOVA analysis of models from linear to 5th degree polynomial

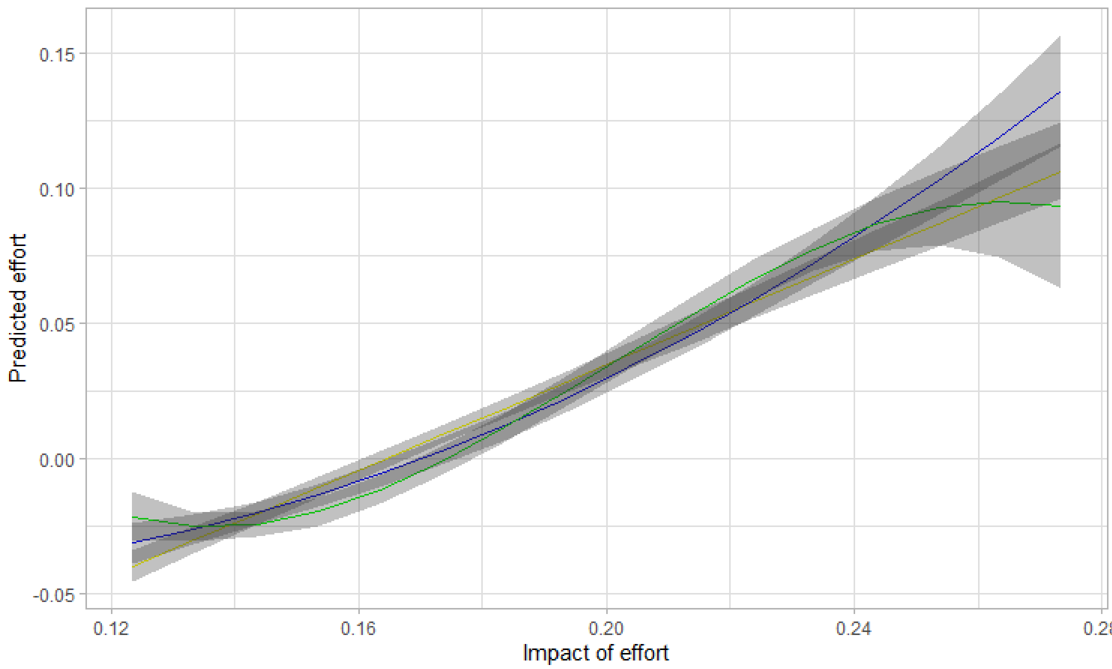


Figure 27: Comparisons of predicted effort: Yellow linear model, Blue- quadratic model, Green- 3rd degree polynomial