Sharing when stranger equals danger: Ridesharing during Covid-19 pandemic

Marc Ivaldi Emil Palikot*

September 16, 2020

Abstract

Using data collected from one of the most popular ridesharing platforms, we illustrate how mobility has changed after the exit from the Covid-19 induced confinement. We measure the impact of the Covid-19 outbreak on the level of mobility and the price of ridesharing. Finally, we show that the pandemic has exacerbated ethnic discrimination. Our results suggest that a decision-maker encouraging the use of ridesharing during the pandemic should account for the impact of the perceived health risks on ridesharing prices and should find ways to ensure fair access.

Keywords: Ridesharing, digital mobility, price discrimination

JEL Classification: L91, R40

*Toulouse School of Economics, University of Toulouse-Capitole. Both authors acknowledge funding from the French National Research Agency (ANR) under the Investments for the Future (Investissements d'avenir) program, grant ANR-17-EURE-0010.

1 Motivation

Pandemic notwithstanding, equitable and cost-effective transportation systems are indispensable to the functioning of modern economies. Meanwhile, in recent months, public-health experts have repeatedly called to avoid crowded and enclosed spaces, both typically associated with public transportation.[1] In fact, Harris (2020) argues that New York's subway was "a major disseminator if not the principal transmission vehicle" in the city's Covid-19 outbreak.[2]

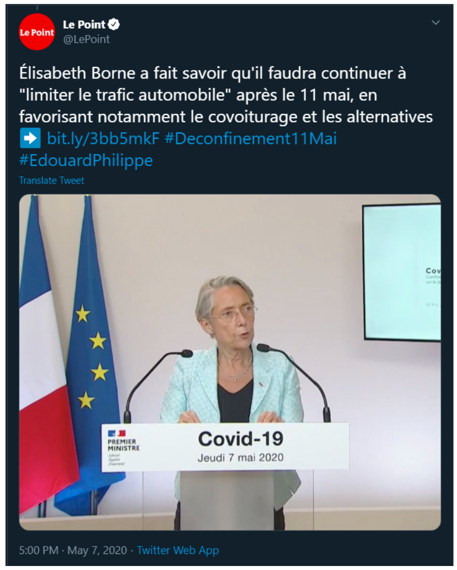
In France, the government encourages ridesharing as the country emerges from the Covid-19 induced lock-downs. Ridesharing, contrasted with traditional means of public transportation, involves a lower number of people traveling together and therefore, might be a way to keep the transmission rates of the virus at low levels, without restricting mobility. It is unclear whether the government's oral encouragement is sufficient to convince people to get into a stranger's car, especially during a pandemic. Here, we look into what happens to the most popular ridesharing service in France as the country lifts shelter-at-home orders.

The ridesharing platform that we study, called BlaBlaCar, is a marketplace for city-to-city rides, where drivers are generally non-professionals. The platform matches a driver, who decides the route and the price, with passengers. In this way, the service allows the drivers to cover part of their costs and provides cheap transportation for passengers. The flexibility in terms of destinations, departure times, and prices might turn out to be of particular value in the uncertain times of re-opening. In this paper, we use observational data to highlight some developments on the platform over the last two months.

[1] See https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/prevention.html

[2] The final jury on the role of public transportation in the spread of Covid-19 is still out. For example, a recent study by Santé Publique France (2020) shows that none out of 150 Covid-19 clusters identified in France between the 9th of May and 3rd of June were related to public transportation. Furthermore, methodological concerns over the results of Harris (2020) were raised (https://pedestrianobservations.com/2020/04/15/the-subway-is-probably-not-why-new-york-is-a-disaster-zone and https://marketurbanism.com/2020/04/19/automobiles-seeded-the-massive-coronavirus-epidemic-in-new-york-city).

Figure 1: French minister of ecological transition and solidarity encouraging use of ridesharing.



Source: Twitter Le Point account

First, we show rapid growth in the number of trips after the shelter-at-home orders were lifted. This increase happened mostly in regions less affected by the Covid-19 outbreak, in terms of cases, hospitalizations, and deaths. Second, we document a substantial dispersion in prices set by drivers. We also provide evidence suggesting a price premium associated with traveling in this period. The prices are particularly high in the regions most affected by the Covid-19 outbreak. Such an increase in prices could result from a higher perception of health risks associated with traveling with a stranger. We check the robustness of this result by exploring the errors of daily reports in Covid-19 cases and by excluding regions following different re-opening protocol.

Finally, studying reviews left by past passengers we show the ethnic composition of cars and how it changed during the pandemic. In particular, we first, document an increase in the number of passengers from ethnic minorities, and second, we show some evidence suggesting that drivers from the ethnic majority that review travel requests manually are less likely to accept a minority passenger during the Covid-19 outbreak than before it.[3] To obtain this result we exploit the fact that for a large subset of drivers in our data we observe

[3]McLaren (2020) provides evidence of a strong positive correlation between total Covid-19 cases and the share of a minority population in a county. Interestingly, the difference can be explained by the use of public transport.

all the reviews, both written by passengers traveling with these drivers before the Covid-19 outbreak and during it.

This paper relates to a quickly growing economics literature on the Covid-19 outbreak; Brodeur et al. (2020) provides an overview of this literature. Several papers focused on the inequality of the impact of the pandemic across racial or ethnic lines. In a related paper, Bartos et al. (2020), using a representative sample from the Czech Republic, shows a magnified hostility against foreigners, especially towards those from Asia. At the same time, they show no change in attitude towards domestic out-groups (migrants and minorities).

The rest of the article is organized as follows: in section two, we discuss the timing of lock-down orders and the re-opening in France. We present data that we use to measure the geographical differences in the severity of the Covid-19 outbreak. Section three contains further details on BlaBlaCar, discussion of the data collection process, and finally, we present some descriptive statistics. In section four, we present the main descriptive evidence on the number of trips offered and the prices. Section five focuses on ethnic discrimination, and finally, section six concludes.

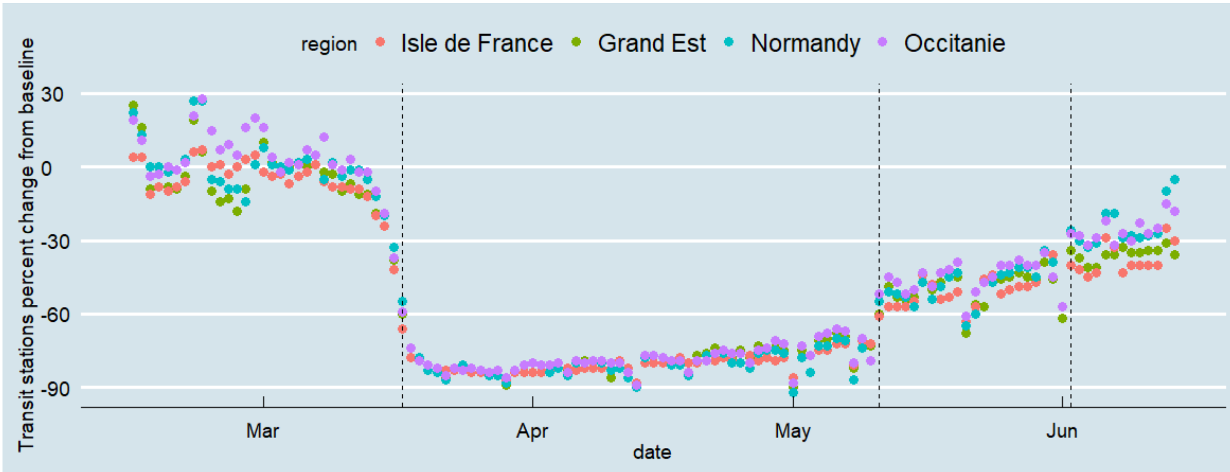
2 Covid-19 outbreak, lock-down, and re-opening in France

France introduced a shelter-at-home order on the 17th of March. Mobility was restricted to essential travels and public transportation was substantially limited. These measures were heavily policed. The first stage of re-opening started on the 11th of May; the restriction to essential travels was lifted. However, the non-essential trips were limited to 100 kilometers from the address of residence. The country was divided into *green* and *red* zones. The *red* zone consisted of departments that were most heavily impacted by Covid-19 and some of the measures were enforced longer.[4] Finally, on the 2nd of June, the limit to 100 kilometers was lifted.

Figure 2 shows the change in mobility trends in places related to public transport hubs such as subway, bus, and train stations from Google's Covid-19 Community Mobility report. The three vertical lines are on the: 17th of March (lock-down), 11th of May (1st stage opening), and 2nd of June (2nd stage opening). From this figure, we can see a sharp drop in mobility following the introduction of the shelter-at-home measures and a gradual increase as the country started the re-opening.

[4] Department (from French département) is the administrative unit of France. It is the middle of the three levels of the regional government, between the regions and the communes.

Figure 2: Mobility in France during Covid19 outbreak



Note: Change in mobility in public transport hubs from baseline in selected regions of France. Vertical lines from left: 17th of March (lock-down), 11th of May (1st stage opening), and 2nd of June (2nd stage opening). Source: https://www.google.com/covid19/mobility/

Grand Est and *Ile de France* are regions most impacted by the Covid-19 outbreak, while *Normandy* and *Occitanie* were amongst the less influenced regions. We can see towards the end of the graph (late May and June) that regions less affected were approaching the baseline faster than *Grand Est* and *Ile de France.*

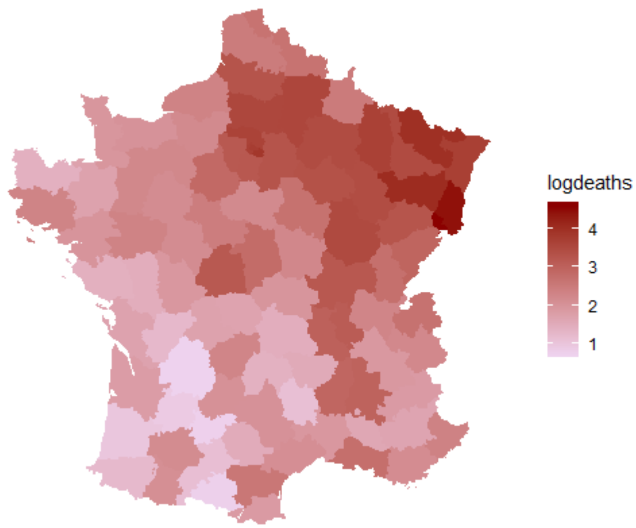
Covid-19 outbreak and related measures: We use four variables, at the daily frequency, to proxy the severity of the Covid-19 outbreak per department. First, we have the number of people that tested positive for the virus per thousand people in the department - (*positive*). We also include the number of tests carried out in the department (*tested*). Second, we use *hosp* and *deaths* which are the number of hospitalizations per thousand inhabitants in the department and the number of deaths per thousand. The source of this data are daily reports available at *https://www.data.gouv.fr/*.

In all regressions that follow we use sums of these measures per department (sums across time, ie, day); thus, we exploit the differences across departments.[5] Figure 3 presents the geographical variation in the number of Covid-19 related deaths in France.

Additionally, we define a binary variable *red zone*, taking the value of one when the observed trip originates from a city in a department designated as heavily impacted by the

[5]These data have also an interesting time dimension, which shows the Covid-19 outbreak in France in time, which we discuss in Appendix A.

Figure 3: Covid-19 outbreak in France



Note: Logarithm of the number of deaths per 100 thousand per department. Source https://www.data.gouv.fr/

Covid-19 outbreak and zero otherwise.

All measures defined above and used in the subsequent analysis refer to the city of departure. Finally, to control for the heterogeneity across departments, we include the share of population above 60 and 75 years per department, and the total population; these variables come from INSEE - the national statistics bureau of France.

3 BlaBlaCar and the data collection

BlaBlaCar is the global leader of city-to-city ridesharing; it is particularly popular in France, where it was established. BlaBlaCar drivers are mostly non-professionals that are looking for a way to cover some costs of day-to-day commutes or longer trips. On BlaBlaCar, the driver sets the price; the passenger observes available drivers and sends a booking request. Some drivers are using automatic acceptance, others reserve the option to manually review the requests.

BlaBlaCar operates at the long tail of transportation; it is most popular on routes where public transportation is of low quality or at times of the day when it is hard to find a ride. The French government has been actively promoting ridesharing. In France, there are over

2000 zones for ridesharing (*aires de covoiturage*), which are spots close to key transportation hubs and highway entrances that are dedicated drop-off and pick-up locations.

We have started collecting data from BlaBlaCar on the 8th of May, that is during the lock-down in France, and three days before the first stage of reopening and continued until the 12th of July. It's worth pointing out that until the 26th of June BlaBlaCar has restricted the number of seats a driver can offer to one.

The main source of our data is the BlaBlaCar website and it has been collected using the BlaBlaCar's API. Some additional information that is specific to drivers and not available on the API has been collected using a web-scraper. Each day we were looking for all trips that depart from all cities that have a *Hôtel de préfecture*, which means that they can be considered as a capital of the administrative area called department, or in its vicinity. There are in total 102 departments; we have not collected any trips in oversea departments, and we have collected but did not include in the present analysis trips departing from Corsica (there were 76 of them during the sample period).[6]

Table 1 provides summary statistics of main variables.

Table 1: Summary statistics of selected variables

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
price	239,934	16.153	12.040	1.115	6.691	22.304	221.928
distance	239,934	269.328	226.425	0	103	380	1,567
price per km	239,923	0.115	0.302	0.001	0.034	0.124	58.549
average rating	52,375	4.409	1.105	0.000	4.500	4.800	5.000
number of ratings	52,375	42.420	62.029	0.000	7.000	51.000	778.000
red zone	239,934	0.265	0.441	0	0	1	1
hosp	239,934	34.185	27.274	2.852	14.678	50.302	140.141
deaths	239,934	15.766	15.525	0.222	6.102	21.012	101.031
share 60+	239,934	0.284	0.045	0.097	0.253	0.322	0.393
share 75+	239,934	0.103	0.020	0.021	0.089	0.116	0.149
positive	239,934	0.956	1.056	0.002	0.146	1.458	3.790
tested	239,934	0.013	0.013	0.00004	0.004	0.017	0.066
rating	1,889,892	4.653	0.602	1	4	5	5
reviewer minority	1,889,892	0.089	0.284	0	0	0	1
reviewer male	1,889,892	0.489	0.500	0	0	1	1
driver minority	52,375	0.096	0.294	0	0	0	1
driver male	52,375	0.698	0.459	0	0	1	1

[6] The oversea departments are: Guadeloupe, Martinique, the Guianas, Réunion, and Mayotte.

In the data collected from BlaBlaCar's API, we observe the price set by the driver and the commission added by BlaBlaCar. In Table 1 and in the subsequent analysis, we use the price paid by the passenger, which is the price that includes added commission. *Distance* is measured in kilometers (km). *Number of ratings* is the number of reviews (stars from 1 to 5) available on profiles of drivers. *Average rating* is the average of these ratings. Car categories are assigned by BlaBlaCar and include *Berline*, for example Ford Fiesta, Renault Clio, *Van*, e.g., Peugeot 5008, Ford C-MAX, *Break* e.g., Peugeot 508, Audi A3, *Tourism* e.g. Ford Focus, Citroen Picasso, and the last category is *Cabriolet*. We also observe whether a driver decided to signal that they allow smoking and pets and whether the booking is automatic or the driver reviews the requests.

Finally, we used the names available on drivers' profiles to establish their gender and whether the name is either of Arabic or African origin suggesting that the driver is of an ethnic minority. Additionally, we used the names associated with reviews left by the past passengers (both before the pandemic and during) to establish the gender and the ethnicity of passengers. In this way, we also have some information about the driver-passenger pairs before and during the pandemic.

There is a number of missing observations because our data collection routine collects different variables on a different speed. We have collected data on all the trips available from the API; however, drivers' characteristics that were not available from the API (for example ratings and reviews) and were scrapped from the website, are available only for some drivers.

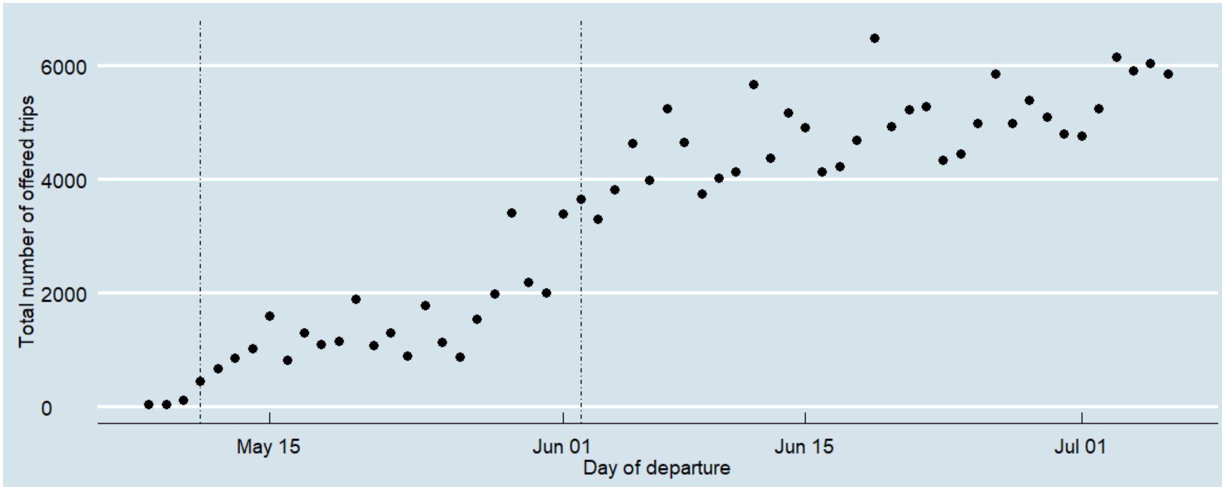
4 Descriptive evidence

In this section, we explore the variation across departments in the severity of the Covid-19 outbreak to study its impact on the number of trips offered on BlaBlaCar and on the prices for these trips.

4.1 Rapid increase in the number of trips post-reopening

With shelter-at-home orders in place and mobility reduced to essential workers, the activity on the BlaBlaCar platform was minimal. As the country started to ease these restrictions, the number of drivers offering trips on BlaBlaCar started to slowly take off. Figure 4 shows the total number of trips per day in the sample period. The two vertical lines indicate the key moments of the de-confinement.

Figure 4: Growth in the number of trips



Note: First vertical line 11th of May is the first stage of easing of restrictions, the second one is the 2nd of June

First, some restrictions were lifted on the 11th of May. Prior to this date, only essential trips were allowed and this has been strictly enforced. After the 11th of May, the non-essential trips up to 100 kilometers from the place of residence were permitted. Next, the line on the 2nd of June shows the second stage of easing of restrictions when traveling in the country became mostly unrestricted. During the period between the 11th of May and the 2nd of June, some departments were indicated as *red*, referring to their higher level of virus circulation. Restrictions in these areas were stricter and enforced more severely.

From Figure 4, we observe that the number trips increased very quickly from the total of 36 trips per day on the 8th of May to around 5000 trips daily in the second part of June and beginning of July.[7]

In the second step, we explore the geographical variation in this growth. To do so, we sum the number of trips by the department and the day of departure and regress these sums on different measures of severity of the virus outbreak. We use four measures: *positive* for the number of people tested positive per thousand, *hosp*, for the number of hospitalization per thousand, *deaths* for the number of deaths per thousand, and *red zone* which is the binary variable taking the value one when the ride departs from a department in the red zone and zero otherwise. The results are presented in Table 2. All regressions have day fixed effects. The dependent variable is in logarithm and we use the OLS estimator. Standard errors are

[7]The dates of opening were announced well in advance.

clustered at the department level.

Table 2: The severity of the virus' impact determines the number of trips

	Dependent variable:			
	log total trips per department			
	(1)	(2)	(3)	(4)
positive	-0.14^{**} (0.07)			
hosp		-0.01^{**} (0.002)		
deaths			-0.01^{**} (0.004)	
red zone				-0.44^{***} (0.15)
total pop	0.001^{***} (0.0001)	0.001^{***} (0.0001)	0.001^{***} (0.0001)	0.001^{***} (0.0001)
Constant	-0.13 (0.24)	-0.08 (0.24)	-0.11 (0.24)	-0.16 (0.24)
Observations	5,602	5,602	5,602	5,602
R^2	0.57	0.57	0.57	0.59
Adjusted R^2	0.57	0.57	0.56	0.58
Residual Std. Error (df = 5535)	0.77	0.77	0.77	0.76
F Statistic (df = 66; 5535)	112.10^{***}	113.24^{***}	111.15^{***}	118.29^{***}
Note:				*p<0.1; **p<0.05; ***p<0.01

Note that all our measures of the severity of the Covid-19 outbreak are related to a lower number of trips departing from that department, even after controlling for the total population. The magnitude of this effect is quite substantial. An increase from the 25th percentile to the 75th percentile in the number of positive tests is associated with a decrease in the number of trips by 22%, as indicated by the first regression. Using the regression in column two, hospitalization per thousand, we find that the change from the 25th percentile in the number of hospitalizations per thousand to the 75th percentile is correlated with a 21% decrease in the number of trips departing from the city.

4.2 Prices during the pandemic

BlaBlaCar drivers set prices themselves; they receive a recommendation from BlaBlaCar that the price should be 0.062 EUR per km. At this price, a driver should be able (for an average car) to roughly cover the costs of the trip by selling two seats. However, this suggestion is not prominently displayed and generally, most of the drivers deviate from it.

There are several reasons why prices in the period of a pandemic might differ from "normal" times, for example:

1. Low demand due to the introduced measures: during a substantial part of the period of the analysis, the non-essential movement of people was restricted to less than 100 km.

2. Low supply of transportation alternatives: in the period of the study, severe restrictions in public transportation were in place, and the number of drivers on the BlaBlaCar was much lower than during the pre-pandemic period. Moreover, drivers were allowed to sell only one seat in their cars to promote social distancing.

3. Health risks associated with sharing a ride: higher risk as perceived by passengers can result in a decrease in the demand for seats. On the other hand, drivers might find it more costly (in terms of expected health costs) to have someone in the car, thus they might set a higher price. We conducted a simple text analysis of the comments associated with the listings and we find that 20% of drivers mention "masks" in their description. This suggests that drivers consider health hazards.

The direction of the change in prices, as compared to pre-pandemic levels would indicate which of the above-mentioned effects dominate. Unfortunately, we do not have access to comparable prices from before the Covid-19 outbreak. We will, however, use the geographical variation in the severity of Covid-19 measures to investigate, which of these forces are likely to prevail. Additionally, we provide some comparison with Lambin and Palikot (2019) (LP herein), which has a large dataset encompassing listings posted in 2017 and 2018.

Figure 5 presents a distribution of prices per kilometer, the green line is the mean. We provide two points of reference: the red vertical line is the price suggested by BlaBlaCar, and the orange line is a mean price from LP. The mean price per kilometer is substantially higher both than the price recommended by the BlaBlaCar and the price documented in LP. This suggests that the low supply and health risk as perceived by the drivers are the dominant factors.

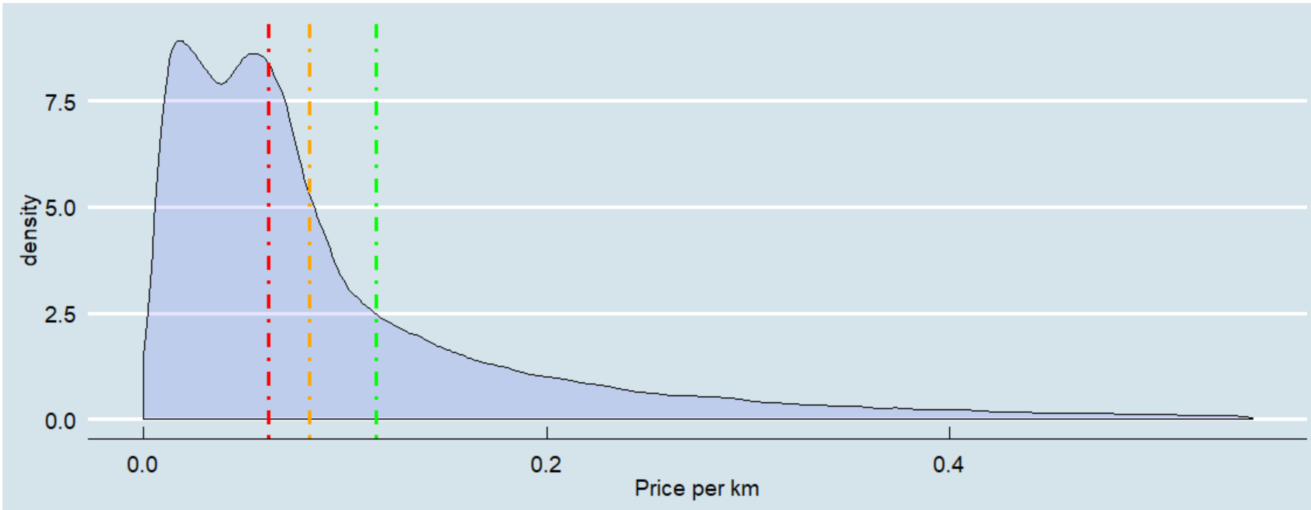
However, a closer inspection of this distribution shows that there is substantial dispersion in prices. The price on the 25th percentile is 0.03 EUR per km while the price on the 75th percentile is 0.12 EUR per km. Furthermore, the median price is 0.065; thus, most of the drivers set prices that are around the price suggested by the BlaBlaCar. The Gini coefficient of prices in our dataset is 0.46, while in the LP it is 0.11.[8] A substantially higher Gini coefficient provides further evidence of the high price dispersion. Finally, the coefficient of

[8]The Gini coefficient is defined as

$$G = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n}|x_i - x_j|}{2n^2\bar{x}},$$

where x are observed prices n is the number of observations and $\bar{x}$ is the mean price. To calculate the Gini coefficient and the coefficient of variation we have used only a subset of our prices associated with trips that depart or arrive to Paris in order to make the sample more comparable to the LP sample. Including the entire dataset does not alter the conclusion.

Figure 5: Distribution of prices per km



Note: Red vertical line - price recommended by BlaBlaCar; orange vertical line - mean price in LP; green vertical line - mean in the current sample

variation (the ratio of the standard deviation to mean) is significantly higher too: in the current dataset, it amounts to 1.11, while in LP to 0.83. Thus, it has increased by 34%. We interpret this as a 34% increase in perceived risk or in the uncertainty of offering the service, for which drivers require compensation, so they set higher prices.

To shed some light on the factors that contribute to the differences in prices, we provide regressions controlling for the severity of the Covid-19 outbreak in the department from which the ride departs. Results are in Table 3. Here we also include controls for the driver level of reputation (the number of ratings and the average rating); additionally, we include whether the driver has a manual acceptance of booking requests, i.e. if she reserves the right to accept or reject the passenger, driver's gender, smoking and pet policy, and the category of the car. Each of the regressions control for day fixed effects. We use the OLS estimator and cluster the standard errors at the level of department.[9]

We observe that all measures of the impact of the virus are associated with higher prices per kilometer. The impact is highly statistically significant. Furthermore, the economic significance of it is also substantial. Around the mean, an increase in the number of positive cases per thousand by one unit is associated with an increase in prices by 6.5% while, an increase in the rate of hospitalization by a unit leads to prices higher by 0.25%.

[9] In Appendix B, we show results of similar regressions that include only controls collected with BlaBlaCar API, which results in a substantially increased sample.

Table 3: The impact of the Covid-19 circulation on the price per kilometer

	Dependent variable:			
	log(price km)			
	(1)	(2)	(3)	(4)
distance	-0.003^{***} (0.00004)	-0.003^{***} (0.00005)	-0.003^{***} (0.00004)	-0.003^{***} (0.00005)
positive	0.065^{***} (0.012)			
hosp		0.003^{***} (0.001)		
deaths			0.003^{***} (0.001)	
red zone				0.067^{**} (0.031)
total pop				0.0001^{***} (0.00003)
manual booking	0.192^{***} (0.013)	0.192^{***} (0.013)	0.193^{***} (0.013)	0.198^{***} (0.012)
driver male	0.162^{***} (0.013)	0.162^{***} (0.012)	0.165^{***} (0.012)	0.163^{***} (0.012)
average rating	0.024^{***} (0.007)	0.023^{***} (0.007)	0.023^{***} (0.007)	0.024^{***} (0.007)
number ratings	-0.0004^{***} (0.0001)	-0.0004^{***} (0.0001)	-0.0004^{***} (0.0001)	-0.0004^{***} (0.0001)
smoking	0.082^{***} (0.010)	0.085^{***} (0.011)	0.084^{***} (0.011)	0.083^{***} (0.010)
pets	-0.019 (0.013)	-0.021^{*} (0.013)	-0.022^{*} (0.013)	-0.019 (0.012)
Constant	-1.631^{***} (0.134)	-1.672^{***} (0.147)	-1.614^{***} (0.141)	-1.753^{***} (0.161)
Car Category	x	x	x	x
Day FE	x	x	x	x
Observations	37,873	37,873	37,873	37,873
R^2	0.456	0.456	0.454	0.457
Adjusted R^2	0.455	0.455	0.453	0.456
Residual Std. Error	0.759 (df = 37796)	0.759 (df = 37796)	0.760 (df = 37796)	0.759 (df = 37795)
F Statistic	417.403^{***} (df = 76; 37796)	417.131^{***} (df = 76; 37796)	414.329^{***} (df = 76; 37796)	413.186^{***} (df = 77; 37795)
Note:				*p<0.1; **p<0.05; ***p<0.01

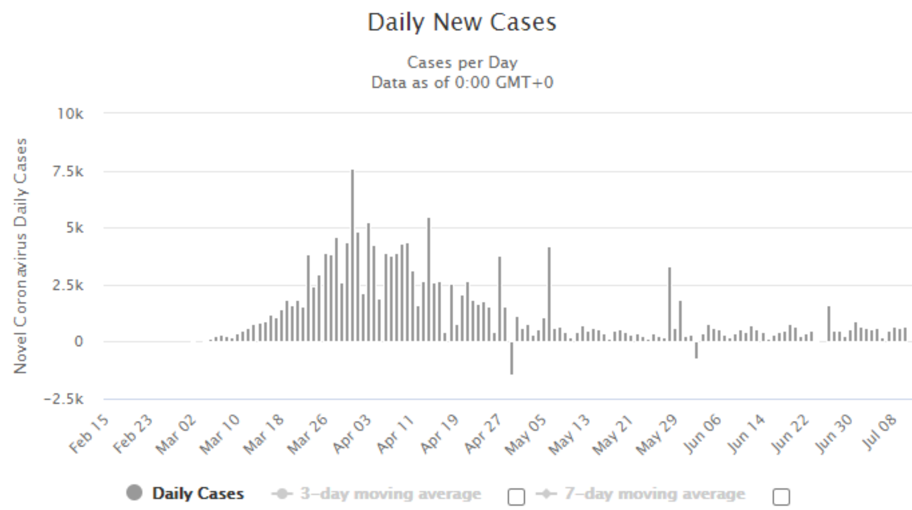
4.2.1 Robustness

Our findings are based on observational data, which does not allow us to claim with full confidence a causal nature of the identified patterns. However, to reinforce the conclusion that the increase in perceived health risk results in an increase in prices, we provide two robustness checks. First, we exploit misreports in daily counts of Covid-19 cases as a source of exogenous variation in the perception of such risk. Second, we exclude *red* regions, due to differences in the enforcement of shelter-at-home orders.

Misreports of daily Covid-19 counts: Daily reports of the Covid-19 cases exhibit sudden jumps; we show this in Figure 6. These day to day jumps, or discontinuities, might be due to several reasons such as, for example, mistakes in earlier reporting or a large batch of tests carried out on a given day. It is, however, unlikely that they are due to a sudden change in the number of infections (especially the negative shocks).

One might worry that our claim of the risk premium is due to confounders. That is, there might be omitted (from regressions) region-specific characteristics that are both correlated with the number of Covid-19 cases and the prices. Thus, the positive relationship that we showcase is spurious. If there are, indeed, important variables that we have omitted from our analysis, they are probably unrelated to the errors in reporting of Covid-19 cases. Hence,

Figure 6: Daily reports of Covid-19 cases in France



Source: https://www.worldometers.info/coronavirus/country/france/

if drivers and passengers use the daily reports in the number of Covid-19 cases, to evaluate the risk of traveling by ride-sharing, we can exploit the errors in these reports as a source of exogenous variation.

In the time-window of our dataset, there are four shocks: three of them are positive (28th and 30th of May, and 26th of June) and one is negative (2nd of June). Drivers' perception of the health hazard just before the positive (negative) jump should be lower (higher) than just after it. Indeed, the prices in our dataset seem to react to such discrete changes. The prices one day before the positive jump are 6% lower than one day after it. The negative shock has the opposite impact on the perception of risk. So when we include both types of shocks, we compare the prices after the negative shock and before the positive one, with prices before the negative shock and after the positive one. Such comparison leads to a change in prices by 3.4%.

To control for potential changes in drivers' characteristics before and after such shocks, region-specific effects, and the time trend, we carry out regression analysis. To do so, we use the sub-sample that includes only trips that depart one day before and one day after these discontinuities. Results are in Table 4. In the second regression, we control for time trend and in the third one, additionally, for department fixed effects. We are interested in the variable *shock*, which takes the value of 1 after positive shocks (31st of May and 27th of June) and before the negative one (1st of June), and zero otherwise (27th of May, 3rd of June, and 25th of June).[10] As before, standard errors are clustered at the department level.

Table 4: The impact of the Covid-19 reporting shocks on the price per kilometer

	Dependent variable:		
	log(price km)		
	(1)	(2)	(3)
shock	0.080*** (0.028)	0.084*** (0.029)	0.084*** (0.029)
distance	−0.003*** (0.0001)	−0.003*** (0.0001)	−0.003*** (0.0001)
manual booking	0.212*** (0.023)	0.211*** (0.023)	0.199*** (0.022)
driver male	0.197*** (0.021)	0.200*** (0.021)	0.193*** (0.022)
average rating	0.008 (0.013)	0.007 (0.013)	0.007 (0.013)
number ratings	−0.0001 (0.0002)	−0.0001 (0.0002)	0.00003 (0.0002)
smoking	0.062*** (0.019)	0.063*** (0.019)	0.065*** (0.019)
pets	−0.024 (0.023)	−0.022 (0.023)	−0.024 (0.024)
Constant	−2.291*** (0.079)	−48.165 (32.095)	−53.793* (31.669)
Car Category	x	x	x
Time Trend		x	x
Department FE			x
Observations	6,120	6,120	6,120
R^2	0.485	0.485	0.511
Adjusted R^2	0.483	0.484	0.503
Residual Std. Error	0.757 (df = 6105)	0.757 (df = 6104)	0.743 (df = 6014)
F Statistic	409.915*** (df = 14; 6105)	383.100*** (df = 15; 6104)	59.870*** (df = 105; 6014)
Note:			*p<0.1; **p<0.05; ***p<0.01

The increase in price following a sudden change in the number of reported cases is robust to changes in drivers' characteristics, time-trend, and the region fixed effects. To calculate the magnitude we use Kennedy transformation (Kennedy, 1981)

$$\frac{\exp(\hat{\delta})}{exp(0.5 \times \hat{V}(\hat{\delta}))} - 1$$

,where $\hat{\delta}$ is the estimated coefficients and $\hat{V}(\hat{\delta})$ is the variance of it. In this way, we find that the impact of the dummy *shock* is 8.7% (we use the coefficient form the last column). The estimated coefficient does not take into account, the fact that the included shocks are of different magnitude. Nevertheless, we conclude from this that an increase in perception of health risk translates into higher prices.

[10] We have excluded the 29th of June because it's between two positive shocks.

Excluding *red* regions: In the next step, we exclude regions marked as *red*. Indeed, since the level of public transportation restrictions, and the enforcement of the ban on non-essential trips over 100km, might have been stricter in these departments, the prices there would be further increased due to no health hazard risks. The results are in Table 5.

We find that even excluding the regions which were hit the hardest and just exploring the differences in the level of virus circulation within the *green* departments, we find that there is a higher level of prices per kilometer when there are more cases, hospitalizations, and deaths related to Covid-19.

Table 5: The impact of the Covid-19 circulation on the price per kilometer excluding red regions

	Dependent variable:		
	log(price per km)		
	(1)	(2)	(3)
distance	-0.003^{***} (0.0001)	-0.003^{***} (0.0001)	-0.003^{***} (0.0001)
positive	0.039^{***} (0.012)		
hosp		0.002^{**} (0.001)	
deaths			0.005^{*} (0.003)
manual booking	0.204^{***} (0.014)	0.205^{***} (0.014)	0.205^{***} (0.014)
driver male	0.170^{***} (0.013)	0.169^{***} (0.013)	0.169^{***} (0.013)
average rating	0.019^{**} (0.009)	0.019^{**} (0.009)	0.019^{**} (0.009)
number of ratings	-0.001^{***} (0.0001)	-0.001^{***} (0.0001)	-0.001^{***} (0.0001)
smoking	0.073^{***} (0.012)	0.074^{***} (0.012)	0.074^{***} (0.012)
pets	-0.026^{*} (0.015)	-0.027^{*} (0.014)	-0.026^{*} (0.014)
Constant	-1.759^{***} (0.047)	-1.886^{***} (0.074)	-1.865^{***} (0.075)
Car Category	x	x	x
Day FE	x	x	x
Observations	28,161	28,161	28,161
R^2	0.463	0.463	0.463
Adjusted R^2	0.461	0.462	0.461
Residual Std. Error (df = 28086)	0.750	0.750	0.750
F Statistic (df = 74; 28086)	327.015^{***}	327.147^{***}	326.852^{***}
Note:			*p<0.1; **p<0.05; ***p<0.01

To conclude, at this stage, we cannot completely isolate the impact of the increased health hazard on the prices. The three factors mentioned earlier (the demand decrease, the negative shock to supply of other means of transportation, and health risk) are likely all at play.[11] Nevertheless, we can conclude that the health factor plays an important role: higher prices in regions that are more affected by the virus (even in the restricted sample), the

[11]In Appendix C, we try to disentangle the supply and demand effects by studying the number of sold seats.

frequent mentioning of the need to have masks in the shared car, and the impact of daily reports suggest that drivers are taking into account the health risk factor.

5 Impact of the pandemic on ethnic discrimination

In our sample, there are only drivers that used the platform at least once after the 8th of May. However, most of them used the platform before as well; the average number of ratings is 43. Thus, for each driver in our sample, we have reviews that she obtained before and during the pandemic.[12] Furthermore, we also know the names of reviewers which allows us to establish the ethnicity and gender of the passengers. In total, we have 278 thousands of reviews left during the pandemic and 2.35 millions of reviews left before it.[13]

We observe that the share of passengers that are of ethnic minority increases substantially during the pandemic, specifically from 9% to 14%. We document this increase in Figure 7, which shows the share of minority passengers per day, starting from January 2019. We see that the share of minority passengers was rather stable before the 17th of March, during the lock-down the share of minority reviewers was high and volatile (there is also much fewer observations in this period - 30 thousand), after the shelter-at-home orders were lifted the share of minority passengers stayed at a level higher than pre-pandemic.

Furthermore, we observe that the increase in the number of minority passengers occurs both in cars with minority and nonminority drivers. A nonminority driver before the pandemic had on average 8% of minority passengers, and 13% during the pandemic, while minority drivers increased their share of minority passengers from 16% to 23%.[14]

Again, several factors could contribute to a higher share of minority passengers during the lock-down: for example, minorities might be more likely to be essential workers, thus during the lock-down they had to commute to work.[15]

Note that it is also important to recognize that passengers before and during the pandemic have access to a different set of drivers. For example, if, during the pandemic, we would observe a higher share of ethnic minority drivers, a passenger who prefers a majority driver

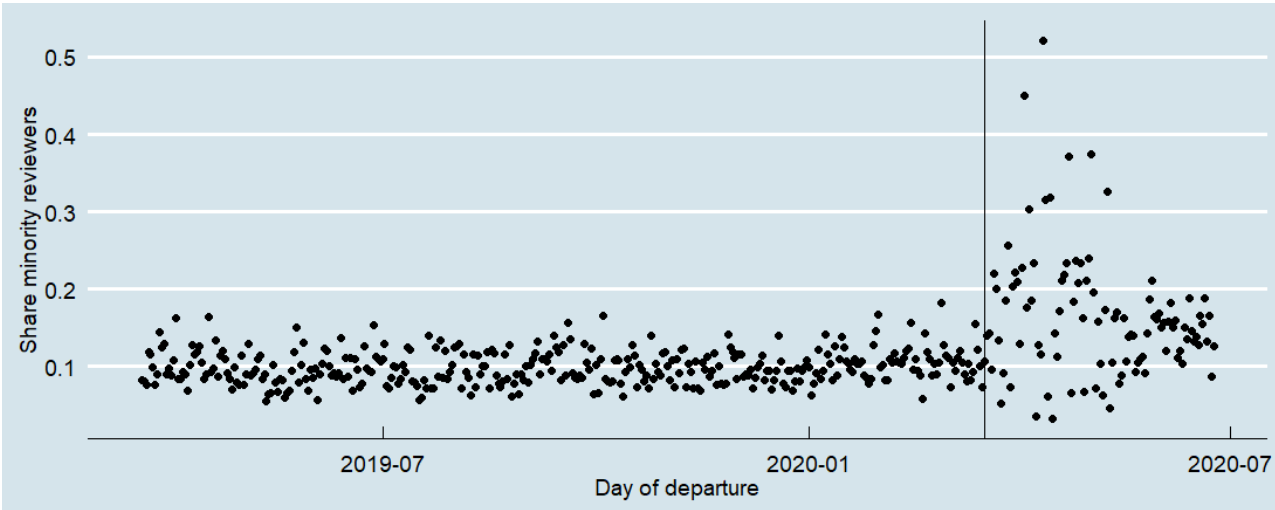
[12]In this section we use the 17th of March as the start date of the pandemic, which was the day in which lock-down has been implemented.

[13]Prior literature used names to establish ethnicity; see for example Rubinstein and Brenner (2014); Laouénan and Rathelot (2020). For the purpose of this section, we assume that the propensity to leave a review is the same during the pandemic and before. Note, that generally, a very high share of passengers leave reviews on BlaBlaCar.

[14]Such horizontal sorting is partly due to the route or the day of the trip. The means computed above do not control for such factors.

[15]Borjas and Cassidy (2020) and Coven and Gupta (2020) provide evidence of this from the United States.

Figure 7: Share of minority passengers per day



Note: The share of reviews written by minority passengers per day. Vertical line March 17.

might find it harder to book a trip during the pandemic then before it. In the current version we do not control for the selection into the sample on ethnicity (and gender) of drivers, but one should recognize that they play a role. (However, we observe that the shares of female and minority drivers are roughly similar to those provided by LP).

Prior literature has shown a positive relationship between greater exposure to health threats and negative attitudes towards out-group members. O'Shea et al. (2020) shows that in US higher exposure to infectious diseases, exacerbates racial prejudice. Do we observe that an increased health hazard, as perceived by the drivers, makes them less eager to accept a minority passenger?

Recall that, on BlaBlaCar, drivers might decide to reserve the right to review a booking request of a passenger - to study her profile for example. An interesting point of comparison is whether drivers that review requests are less likely to take a minority passenger on board during the pandemic. To test for this we carried out two regressions; results are displayed in Table 6.

Column one presents the results of an OLS estimator, while column two is a within panel estimator, where we control for driver fixed effects. The dependent variable takes the value one when a passenger is a minority and zero otherwise; Covid-19 takes the value one for reviews left after the 17th of March and zero otherwise. We are mostly interested in the interaction : *driver nonminority * Covid19 * manual.*

We find that, first, during the Covid-19 there is an indeed higher chance of a passenger

Table 6: Share of minority passengers across minority and non-minority drivers

	Dependent variable:	
	passenger minority	
	OLS	*panel linear*
	(1)	(2)
driver nonminority	−0.083*** (0.001)	
number ratings	0.0001*** (0.00000)	
Covid-19	0.071*** (0.004)	0.059*** (0.004)
manual	−0.022*** (0.001)	−0.001 (0.006)
Covid-19*driver nonminority	−0.019*** (0.004)	−0.014*** (0.005)
Covid-19*manual	0.014** (0.006)	0.007 (0.006)
driver nonminority*manual	0.015*** (0.001)	0.001 (0.006)
driver nonminority* Covid-19 * manual	−0.011* (0.006)	−0.016** (0.007)
Constant	0.161*** (0.001)	
Driver FE		x
Observations	1,955,469	1,955,469
R^2	0.009	0.001
Adjusted R^2	0.009	−0.013
Residual Std. Error	0.286 (df = 1955460)	
F Statistic	2,142.231*** (df = 8; 1955460)	238.664*** (df = 6; 1929823)
Note:		*p<0.1; **p<0.05; ***p<0.01

being a minority. Second, non-minority drivers are less likely to have a minority passenger on board, and it is even less so during the pandemic. Finally, the coefficient of the interaction: *driver nonminority * Covid19 * manual* has a negative sign. Thus, non-minority drivers that use manual acceptance are even less likely to accept a minority passenger during the pandemic.[16]

6 Conclusion

Our results suggest that a decision maker encouraging the use of ride-sharing during or right after the Covid-19 outbreak should take into account the likely increase in prices, particularly in the regions which are already severely affected by the pandemic, and exacerbated ethnic/racial inequality of access.

In our view, the price premium might be in part due to an increase in the perception of health hazards associated with traveling with a stranger in the car. We provide evidence suggesting that even daily reports of the number of cases shape drivers' perception of such

[16] We have also carried out an analogous analysis concerning male versus female passengers. We do not find any substantial differences due to the pandemic.

risks. This result highlights the importance of the provision of accurate and up-to-date information.

Design and comparison of policies that might ensure fair access to the service are out of the scope of this article. There is, however, a rich literature studying discrimination in online markets. For example in the context of ride-sharing, Ge et al. (2016) provides a comparison of different platforms, which suggests that removing names from trip booking has the potential to alleviate the immediate problem. Lambin and Palikot (2019) develops a structural model which also shows that the removal of names can mitigate ethnic discrimination.

Finally, we would like to reiterate that the results obtained in this article are based on observational data and we cannot have full confidence in their causal interpretation. Although we have provided some robustness checks and controlled for a number of other potential effects, further work is needed to fully understand the impact of an increase in risk perception, due to a pandemic, on prices, and on access to the service of (ethnic) minorities.

References

Bartos, V., Bauer, M., Cahlíková, J., and Chytilová, J. (2020). Covid-19 crisis fuels hostility against foreigners.

Berry, S., Levinsohn, J., and Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica: Journal of the Econometric Society*, pages 841–890.

Borjas, G. J. and Cassidy, H. (2020). The adverse effect of the covid-19 labor market shock on immigrant employment. Technical report, National Bureau of Economic Research.

Brodeur, A., Gray, D. M., Islam, A., and Bhuiyan, S. (2020). A literature review of the economics of covid-19.

Coven, J. and Gupta, A. (2020). Disparities in mobility responses to covid-19. Technical report, NYU Stern Working Paper.

Ge, Y., Knittel, C. R., MacKenzie, D., and Zoepf, S. (2016). Racial and gender discrimination in transportation network companies. Technical report, National Bureau of Economic Research.

Harris, J. E. (2020). The subways seeded the massive coronavirus epidemic in new york city. *NBER Working Paper*, (w27021).

Kennedy, P. E. (1981). Estimation with correctly interpreted dummy variables in semilogarithmic equations [the interpretation of dummy variables in semilogarithmic equations]. *American Economic Review*, 71(4):801–801.

Lambin, X. and Palikot, E. (2019). The impact of online reputation on ethnic discrimination. *TSE Mimeo*.

Laouénan, M. and Rathelot, R. (2020). Can information reduce ethnic discrimination? evidence from airbnb.

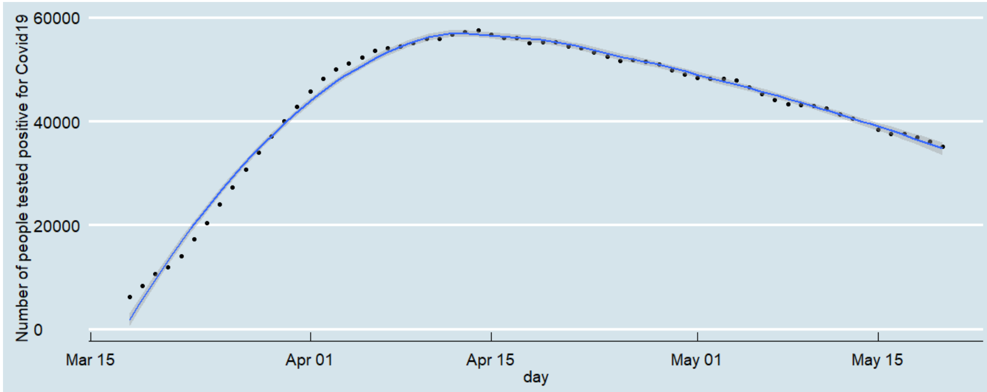
McLaren, J. (2020). Racial disparity in covid-19 deaths: Seeking economic roots with census data. Technical report, National Bureau of Economic Research.

O'Shea, B. A., Watson, D. G., Brown, G. D., and Fincher, C. L. (2020). Infectious disease prevalence, not race exposure, predicts both implicit and explicit racial prejudice across the united states. *Social Psychological and Personality Science*, 11(3):345–355.

Rubinstein, Y. and Brenner, D. (2014). Pride and prejudice: Using ethnic-sounding names and inter-ethnic marriages to identify labour market discrimination. *Review of Economic Studies*, 81(1):389–425.

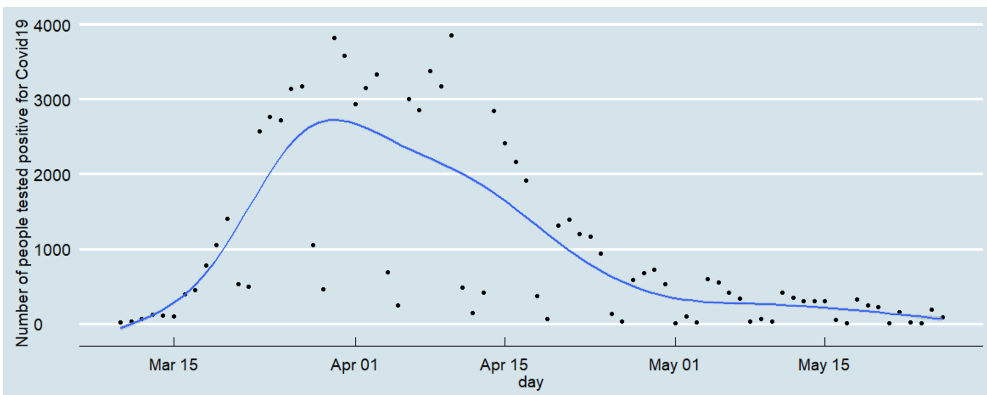
Santé Publique France (2020). Covid-19 point épidémiologique hebdomadaire du 4 juin 2020. Technical report.

Figure 8: Covid-19 outbreak in time



Note: Sum of people hospitalized. Source https://www.data.gouv.fr/

Figure 9: Covid-19 outbreak in time



Note: The number of people tested positive per day. Source https://www.data.gouv.fr/

A Timing of Covid-19 pandemic

In Figure 3 we have presented the geographical variation in the Covid-19 outbreak. However, these data have also an interesting time dimension, which shows the timing of the pandemic. In Figures 8 and 9, we show daily sums of the number of hospitalizations and positive test.

The two measures are closely related. It is also worth noticing that the shelter-at-home orders were introduced just before the height of the pandemic and lasted throughout the peak of it.

B Price regressions with larger sample

In the analysis of section 4.2, we have included driver characteristics collected using a web-scrapper. This has allowed us to account for the impact on prices of these characteristics, but at the same time substantially reduced the sample size.

Table 7: Impact of Covid-19 outbreak on prices

	Dependent variable:			
	log(price per km)			
	(1)	(2)	(3)	(4)
distance	-0.003^{***} (0.00004)	-0.003^{***} (0.00004)	-0.003^{***} (0.00004)	-0.003^{***} (0.00004)
positive	0.068^{***} (0.016)			
hosp		0.003^{***} (0.001)		
deaths			0.004^{***} (0.001)	
red zone				0.088^{***} (0.031)
total pop				0.0001^{***} (0.00004)
booking manual	0.184^{***} (0.007)	0.184^{***} (0.007)	0.184^{***} (0.007)	0.192^{***} (0.007)
Constant	-1.540^{***} (0.153)	-1.604^{***} (0.117)	-1.526^{***} (0.122)	-1.707^{***} (0.134)
Day FE	x	x	x	x
Observations	239,923	239,923	239,923	239,923
R^2	0.465	0.467	0.464	0.470
Adjusted R^2	0.465	0.467	0.464	0.470
Residual Std. Error	0.763 (df = 239807)	0.761 (df = 239807)	0.763 (df = 239807)	0.759 (df = 239806)
F Statistic	$1{,}811.323^{***}$ (df = 115; 239807)	$1{,}826.963^{***}$ (df = 115; 239807)	$1{,}806.699^{***}$ (df = 115; 239807)	$1{,}831.579^{***}$ (df = 116; 239806)
Note:				*p<0.1; **p<0.05; ***p<0.01

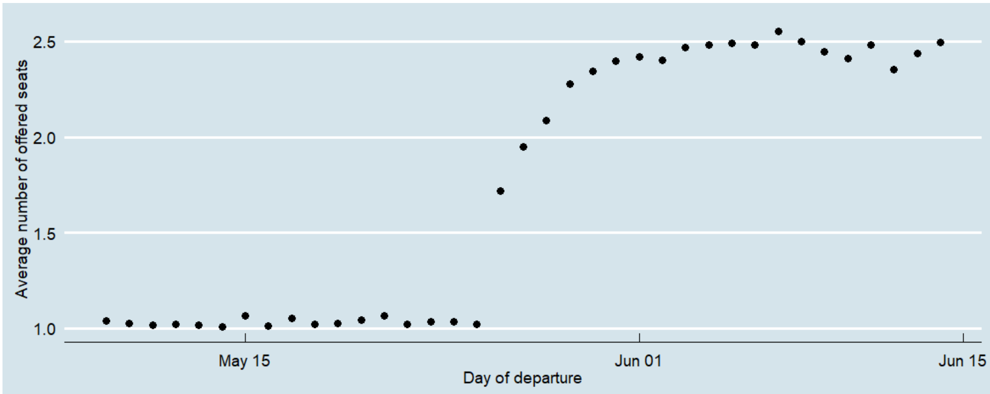
In Table 7, we show the results of regressions with prices as the dependent variable and Covid-19 outbreak measures as explanatory variables. The sample size is increased from 37 thousands to 239 thousands. The conclusion and even the magnitude of the effect are similar to the baseline specification: prices are higher in the departments with more Covid-19 cases.

C Impact of Covid-19 on the number of sold seats

During the height of the Covid-19 pandemic BlaBlaCar has restricted the number of seats a driver can offer on the platform to one. During the de-confinement, starting from the 26th of May, this restriction has been lifted. Figure 10 documents it, by showing the average per day of the number of seats offered by drivers.

Using the part of the dataset after the 26th of May we can study the determinants of the number of sold seats per driver. The variables of key interests include price, measures of the Covid-19 outbreak, and driver characteristics. We are interested in the elasticity of passengers' to changes in price and the impact of the virus. Note, that the availability of other means of transportation might differ depending on the severity of the outbreak in a given department.

Figure 10: Average number of offered seats



Note: Average number of offered seats over the data period; on the 26th of May the constraint to maximum one seat has been lifted.

To address the endogeneity of prices and quantity we introduce instrumental variables. We use sums of characteristics of other drivers (driver male, average rating, number of ratings) and the number of drivers departing from the same city on the same day, and the number of listings on the same route. Such measure of distance in characteristics space influence the markups that drivers can achieve but should not influence the utility of a passenger Berry et al. (1995)).

Note, that we neither control for the number of seats a driver has offered nor for difference in access to other forms of transportation across days and departments. Results are available in Table 8.

The first column in Table 8, shows the first stage regression results. We conclude that the instruments we proposed are statistically significant. Regressions (2) to (4) show that passengers' have negative elasticity to changes in prices. Coefficients associated with the number of the Covid-19 cases and hospitalizations are positive and significant, while the number of deaths is statistically insignificant. Therefore, controlling for the differences in prices, drivers in departments impacted stronger by the Covid-19 are selling more seats.

Table 8: Sold seats as a function of driver characteristics and Covid-19 oubtreak

	Dependent variable:			
	price	sold seats		
	OLS	*instrumental variable*		
	(1)	(2)	(3)	(4)
price		-0.031^{***} (0.008)	-0.020^{***} (0.008)	-0.018^{***} (0.007)
distance	0.005^{***} (0.0003)	0.0002^{***} (0.00005)	0.0001^{***} (0.00005)	0.0001^{***} (0.00004)
positive		0.065^{***} (0.011)		
hosp			0.001^{***} (0.0004)	
deaths				0.001 (0.001)
driver male	2.177^{***} (0.143)	0.133^{***} (0.023)	0.112^{***} (0.022)	0.111^{***} (0.021)
manual booking	2.654^{***} (0.140)	-0.067^{***} (0.026)	-0.094^{***} (0.025)	-0.099^{***} (0.023)
average rating	-0.147^{*} (0.079)	0.104^{***} (0.009)	0.105^{***} (0.008)	0.105^{***} (0.008)
number of ratings	-0.011^{***} (0.001)	0.001^{***} (0.0001)	0.001^{***} (0.0001)	0.001^{***} (0.0001)
smoking	0.994^{***} (0.153)	0.079^{***} (0.018)	0.068^{***} (0.018)	0.065^{***} (0.017)
pets	-0.598^{***} (0.144)	0.010 (0.016)	0.015 (0.016)	0.016 (0.015)
instrument 1	0.307^{***} (0.035)			
instrument 2	-0.041^{**} (0.019)			
instrument 3	-0.001^{**} (0.0003)			
instrument 4	-0.299^{**} (0.116)			
instrument 5	0.011^{***} (0.001)			
Constant	10.892^{***} (0.579)	0.396^{***} (0.101)	0.289^{***} (0.093)	0.295^{***} (0.091)
Car Category	x	x	x	x
Day FE	x	x	x	x
Observations	25,587	25,587	25,587	25,587
R^2	0.082	−0.197	−0.093	−0.076
Adjusted R^2	0.079	−0.201	−0.097	−0.080
Residual Std. Error	10.645 (df = 25500)	1.152 (df = 25503)	1.100 (df = 25503)	1.092 (df = 25503)
F Statistic	26.566^{***} (df = 86; 25500)			
Note:				*p<0.1; **p<0.05; ***p<0.01