The Impact of SNCF Strike on Supply and Demand of Ridesharing: Evidence of BlaBlaCar (Draft)

Timothy Yu-Cheong Yeung[1] and Dianzhuo Zhu[1]

[1]Governance and Regulation Chair, DRM, University Paris-Dauphine, PSL Research University

January 31, 2019

Abstract

This work studies the impact of SNCF strike from April to June 2018 on the demand and supply of the ridesharing market. Using detailed information available through BlaBlaCar official API, we are able to construct the the market supply curve as well as observed market demand curve by transaction records. We find that total transaction payment increased by 37% on an average strike day comparing to an average non-strike day, while market supply increased by only 9%. We then estimate the elasticity of demand by using the relevant price ranges on the observed market demand curve, taken into consideration of the influence of default price on market demand. We conclude that BlaBlaCar generated additional 0.48 - 2 million, or 2.30 - 9.78% of the loss of SNCF, consumer surplus on an average strike day, showing substantial substitutions between SNCF and BlaBlaCar.

1 Introduction

The International Labour Organization (ILO) has recognized that strike action is a fundamental right of workers. Yet, some debates are still going on about whether workers of public services are entitled to the right. According to the ILO, public transportation is not an essential public service and thus its workers should be entitled to the right to strike. As public transportation often occupies a major role in facilitating the normal functioning of an economy, the stakes are so high that labor unions of public transportation are usually very powerful. France experienced one of the worst SNCF (Sociéte Nationale des Chemins de Fer Français / French National Railway Company) strike actions in history from April to June 2018. During the three months, two in every five days a strike was organized, affecting millions of passengers across the nation. When institutionalized mechanisms fail to work, individuals always respond spontaneously. Ridesharing platforms, for example, BlaBlaCar, the successful showcase of start-ups in France, are digital markets for drivers and passengers to match demand and supply. Those platforms are very flexible and thus should be able to absorb additional demands due to the strike. This work investigates this substitution relationship and estimates the impacts of the SNCF strike on the demand and supply of ridesharing. Our result shows that BlaBlaCar absorbed a substantial amount of economic value during the strike period, pointing to the tendency that services provided by public transportation could become more and more easily substituted due to the existence of digital ridesharing platforms.

Through BlaBlaCar's official API, we are able to obtain the details of almost all offers and bookings during the SNCF strike period. The construction of the market supply curve is thus straightforward. As we observe all seat offers, taken or not, along with their prices, we can add up the quantity supplied at each price while assuming that the drivers are willing to accept any offers above the price they quote on the platform. The construction of the market demand curve requires a longer discussion. First, we construct the observed market demand curve by assuming that the maximum willingness to pay coincides with the price a passenger pays. This observed market demand can be considered as the lower bound of the actual market demand curve. The area under the curve is exactly the total payment of all transactions combined. Next, we select the section of the observed market demand curve around the reference price, which is set by BlaBlaCar, while excluding the reference price to estimate the price elasticity of demand. As the supply condition is significantly influenced by BlaBlaCar, most of the offers, and thus bookings, are found at the reference price, producing a spike in both the observed market demand and the market supply curves. However, there are still sufficient amounts of offers and bookings slight above and slight below the reference price, where the influence of the manipulation by BlaBlaCar is smaller. The estimate of the price elasticity of demand is thus plugged into a constant elasticity demand model to estimate the actual market demand curve. By comparing the estimated demand curve and the observed demand curve, we obtain an estimate of consumer surplus generated on the platform of BlaBlaCar.

We find that the impact of strike on the market supply is 9% increase at each price level. The increase implies that either some SNCF passenger became drivers and offered seats on BlaBlaCar, or some new drivers came to the platform to take advantage of the strike. On the other hand, the impact on the observed demand or the payment is roughly 37% increase at each price level. The result confirms our expectation that supply is much less elastic. By comparing the estimated market demand curves on strike and non-strike days, the average increase in consumer surplus on an average strike day generated on the platform of BlaBlaCar is 0.48 - 2 million per day, accounting to 2.30 - 9.78% of the loss of SNCF per strike day.

The paper continues as follows. Next section is a brief literature review on the dynamics between traditional markets and digital platforms, followed by sections on the background of the SNCF strike and on BlaBlaCar. Section 5 discusses the data collection and cleaning, mentioning potential measurement errors, along with some summary statistics. Section 6 and 7 illustrate the constructions of the market supply and the observed market demand curves. Section 8 gives an estimation of the consumer surplus generated by BlaBlaCar and attempts to measure the additional consumer surplus on an average strike day, which could be interpreted as the value salvaged by Blablcar. The last section draws the paper to an conclusion.

2 Literature Review

Transportation network disruptions often occur and have great impacts on the routine and well-being of commuters. Researches on the impact of those interruptions remain city-level case studies, which raise difficulty in the generalization of evidence (Zhu and Levinson, 2012). Strike in the public transportation system is a common category in transportation disruption that forces people to change their commute behavior. Exel and Rietveld (2001) studied 13 strikes in the public transportation sector and concluded that on average only 10-20% of passengers cancelled their trips while others actively sought for alternatives.People may eventually stick to the alternative and form a new habit. A recent paper by Larcom et al. (2017) found that tube/underground strike in 2014 led to lasting changes in commuters' behaviors as they were forced to discover alternatives and some learned that the alternatives are more optimized than their initial commuting routes. Instead of searching for alternatives within the same transportation mode, inter modal switch is also quite common (see Fearnley et al. (2018) for a review of inter modal elasticities), with carpooling or ridesharing as an outside option. However, Exel and Rietveld (2001) showed that despite the evidence on switch to carpooling as a short-term solution, it is often organized either by the authority (as a policy tool) or spontaneously among acquaintances and employees. It is not clear if large-scaled share of ride among strangers is used as a resort and how the long-term effect is.

The rise of digital ridesharing and taxi-alike platforms may now serve as a more efficient substitute during strike and other transportation disruptions, as they can match strangers in a large scale and that they are flexible in adjusting empty seats supplied. What's more, they offer tremendous opportunities for quantitative data analysis on the impact of disruptions, which lacks quantitative evidence. Due to data scale and availability issues, current research on those platforms are concentrated on Uber data of American users, topics include drivers' working behavior under surge pricing (Chen and Sheldon, 2015), gender pay gap (Cook et al., 2018) and racial and gender discrimination (Ge et al., 2016). Some papers attempted to measure the welfare impact of digital platforms. Cohen et al. (2016) exploited the discontinuity in the pricing of Uber to estimate the passenger price elasticity of demand at various points of the demand curve to calculate consumer surplus. Kim et al. (2018) compared Uber with taxi service in New York city and argue that Uber's entry is welfare-enhancing since passengers in broader areas of NYC now have access to taxi or Uber services. Lam and Liu (2017) also used Uber and Lyft data from API as well as taxi data in NYC to estimate the demand and consumer surplus, with a focus on the calculation of surplus on shortened waiting time. In another research using UberPool data in Chicago, Schwieterman and Smith (2018) found 67.6% time reduction using UberPool between neighborhoods and a $.38/minute saving. On the driver side, Chevalier et al. (2018) showed that Uber's flexible working schedule allows drivers to earn twice as much surplus as in non-flexible situations. Evidence suggests that Uber, as a digital platform with flexible on-demand inner-city transportation supply, is welfare enhancing. It increases ridership as a complement of current public transportation system (Hall et al., 2018).

However, when there is disruption in the public transportation system, those digital sharing platforms could also generate extra supply to absorb the excess demand thus act as a substitute of public transportation. The welfare impact would be even more important since they recover the welfare loss caused by disruption should their be no substitute and in a much cheaper, grassroot way. Our paper analyzes rich and unique data extracted from BlaBlaCar during a severe, long-lasting railway stike in France in 2018 and proved the welfare improving impact. SNCF is a state-owned monopoly in the French railway system and represents the majority of market of the long-distance public transportation (Crozet and Guihéry, 2018). The strike has definitely forced people to search for alternatives. As the largest inter-city level long-distance ridesharing platform in Europe, BlaBlaCar has been mentioned in several papers of the sharing economy. Specific papers like Shaheen et al. (2017) investigated the characteristics of passengers and drivers on BlaBlaCar by survey. Farajallah et al. (2019) web scraped data on BlaBlaCar to study the determinants of price and demand for the service and found that more experienced drivers tended to set lower prices. To the best of our knowledge, no economic research paper has been done on the estimation of the economic value brought by BlaBlaCar.

Our paper fills the gap in several aspects. Firstly, it offers quantitative evidence on the behavioral and welfare impact of a severe transportation disruption-SNCF national strike in 2018. Secondly, it is not limited in one city but estimates the impact of entire France thanks to the nationwide representation of BlaBlaCar. Thirdly, it focuses on long-distance trips versus trains rather than daily commute versus subways. Fourthly, it shows the potential of a more-flexible digital sharing platform as substitute to public transportation.

3 Background: SNCF strike and the opportunity for ridesharing

In late March 2018, the railway workers of SNCF (French National Railway Company) decided to start an unprecedented strike. The strike started on 3 April and ends on 28 June. The strike was initiated by the labor union in order to oppose the French government's reform plan of SNCF. The plan includes abolishing "railway worker" status that accompanies lots of advantages comparing to other sectors, privatizing the railway sector to promote competition, closing some unprofitable regional train lines, and letting SNCF itself, but not the state, pays its debts. As for the railway workers, abolishing their status would be the most disadvantageous reform. This status was created in the early 20^{th} century to compensate the difficult working conditions of railway employees at that time. The bonus included lifelong job guarantee, more holidays than average workers, higher retirement pension, lower house rent, and free tickets for themselves and even their close relatives (parents, grandparents,kids and partners).

One special part of the strike is the schedule. The labor unions decided to go on strike every two days out of five from April to June, no matter if the scheduled strike days happen to be weekends or national holidays. As the strike days were announced well before, people could anticipate the inconvenience and make other plans if possible, though SNCF announced the expected train schedules of a strike day only in the morning. Appendix A shows the calendar of strike dates in April, May and June.

The strike became a breaking news in the subsequent few weeks, and impacted on almost every person living in France. Competitors in the transportation sector found their chance to propose solutions and to capture the unserved market. Ridesharing sector was a direct beneficiary. The transportation authority (Île-de-France Mobilités) even formed partnership with 8 inner-city ridesharing platforms and set all ridesharing trips inside Île-de-France free of charge on strike days, including one platform owned by SNCF itself (IDVROOM). For inter-city long-distance trips, where the impact was expected to be the most severe, BlaBlaCar, the most important long-distance ridesharing platform in France, would be the natural substitute for lots of passengers.

4 Background information: Introduction of BlaBlaCar

BlaBlaCar is the largest inter-city ridesharing platform in Europe. It is initially called Covoiturage.fr, created by Vincent Caron in 2004. The domain was then bought by Frédéric Mazzella in 2006, who eventually changed the name into BlaBlaCar.fr in 2013. Besides France, BlaBlaCar is also operating in 21 other countries, all in Europe except Mexico, Brazil and India.

The business model of BlaBlaCar is based on its online platform. Passengers can filter by departure place, destination and departure date to search for the most suitable ridesharing driver. The interface then shows all the filtered trips that drivers have proposed, ranked by an algorithm taking into account driver experience, departure time, available seats, price, trip departure and destination matching, etc. Passengers can view a snapshot of each proposed ride, and then decide which one to click in. Once they enter the individual ride page, they could see more information. They could even click into driver's personal page and learn more about the driver. See Appendix B for more information.

Drivers could not actively search passengers on the platform (there is no "ridesharing request"). BlaBlaCar neither proposes automatic matching like Uber. Drivers can only screen passengers after the latter send requests. However, they can post their trips and have the freedom to set prices. When a driver sets the price level of a trip, the system will propose her a default reference price, but she can choose to increase or decrease the price level. However, the system also sets the highest and the lowest possible price. Drivers could not set beyond the range. The final price that passengers see on the searching page is the price set by drivers plus commission set by BlaBlaCar. Commission levels climb ladders as trip price increases. Appendix C shows how BlaBlaCar sets trip commission.

5 Data

5.1 Data Collection and cleaning

Our data is comprised of three sources: BlaBlaCar's API, BlaBlaCar's website and SNCF's press documents. All information collected is publicly available. However, BlaBlaCar's API will only keep historical data for a short period, so that the data we have for this paper could no longer be found. BlaBlaCar's open API offers the essential part of the dataset. [1] It contains almost all the information of a trip, which could already be seen by everyone, registered to the website or not. Crucial variables are departure and arrival cities, departure date and time, price proposed by driver, commission level, price seen by the passenger, total seats and booked seats. However, no personal information such as driver's name or age is available on API, even though they can also be seen on the page of the trip. For the purpose of this paper, we consider each driver and passenger as independent – though some of them participated several times in the dataset – and we do not need personal information for the analysis.

5.1.1 Protocol of API scraping and choice of routes

The scraping started on 1 April, 2018, two weeks after SNCF's announcement of strike and two days before the strike starts (3 April). Scraping ends on 3 August, 2018. We extended the scraping period one month after the end of the strike because a calm period could serve well as a control group. For the sake of conceiving the protocol and programming, we could not start scraping right after the announcement of strike. As the API keeps some historical trips, we could still include

[1] https://dev.BlaBlaCar.com/

some pre-strike observations in our data base, from 20 March till 31 March. However, once we pass the departure date of the trip for one or two days, only the unbooked trips will remain on the API. Our pre-strike observations are biased because we can no longer know the information of trips that are partially or fully booked.

From 1 April to 3 August, we run the scraping program each day at the end of the working day, typically from 18:00 to 19:00. The choice of scraping frequency is a balance of accuracy and convenience. Scraping more than once per day will include more observations and track changes more precisely, but will soon explode the size of data and exceed the daily upper limit of API queries. Scraping once several days may lose track on important changes that occur on one specific day and confound the effect with idiosyncrasies of the day such as strike or non-strike, weekday or weekend. Since most of the trips depart before 19 p.m., scraping at the end of the day increases the chance of documenting the end status of the trip after its departure.

We also limited our daily scraping to a reasonable level, not only because that it is nearly impossible to exhaust all the trips on the website, but also because that too large dataset will cost too much time spent on scraping each day and will slow down the data analysis process.[2] Following the practice of **?**, we also pre-defined routes that are representative and are the most likely to be impacted by the railway strike.

In the end, 41 round trips (82 routes) are selected. These routes include major French cities (and their suburb areas) and some second-tier cities (and their suburb areas), as shown by Figure 1.[3] We also take into consideration the balance of geographic representation so that each part of the hexagon has represented cities. Routes can be considered of three categories. The first one is between two major cities. We picked seven cities as major cities: Paris, Lyon, Marseille, Bordeaux, Toulouse, Nantes, Strasbourg, balancing the size of the city and the geographic representation. The second category is between a major city to a second-tier city close to that major city. The third category is between two second-tier cities that are close to each other.[4] Appendix D lists the 82 routes and the categories they belong to.

After limiting routes, we limited the length of tracking of trips. It is interesting to start tracking trips several days before its departure in order to learn how booking changes. However, tracing back too many days before departure will also increase burden of scraping for little efficiency. In the end, we decided to trace 14 days beforehand. At scraping date n, we scraped all trips that depart from day n-1 to day n+14. If a trip was published and had bookings more than 14 days before departure date, we will lose track on the first booking date, but such cases are very rare. Graph 2 shows that most trips are published within 3 days before departure. The reason to scrap trips that just finished one day before is that some of those trips may not have been wiped out from API yet, and we might better catch the latest booking status.[5]

5.1.2 Supplementary information from BlaBlaCar and SNCF

Apart from the information available from API, we completed our dataset with supplementary information from BlaBlaCar and SNCF. Even though drivers can propose price themselves, they are not totally free to choose. For each route, when drivers enter the page of price settling, a

[2] The way of scraping via API is by sending queries that contain selection criteria. In the case of BlaBlaCar, we can set selection criteria on various variables like departure and arrival cities, date of departure, etc. By setting a selection criterion, we are restricting ourselves to a subset of the entire data. Also, we are not scraping in permanence. There must be some trips that have appeared but then disappeared in between two scraping sessions.

[3] In 2018, the ten largest cities (including suburb areas) in France are : Paris, Lyon, Marseille-Aix-en-Provence, Toulouse, Bordeaux, Lille, Nice, Nantes, Strasbourg and Rennes. They are all included in our search.

[4] We do not include trips of two cities that are far from each other. For example, we scrap trips from Lille to Paris and from Paris to Lyon, but we will not scrap trips from Lille to Lyon even though Lille and Lyon are both in the list. A trip from Lille to Lyon will definitely pass by Paris, and drivers may well add a correspondence in Paris, which may cause repeated scraping if we scrap Lille-Lyon. However, our scraping logic will split a complete trip of Lille-Paris-Lyon into two observations: Lille-Paris and Paris-Lyon. Since BlaBlaCar asks drivers independent price for each subsection of the entire trip, we can well treat them as two independent observations.

[5] The circumstances at which API wipes out trips are not very clear to us. Based on observation of data, we made some inferences, which will be presented in details in the data cleaning section.

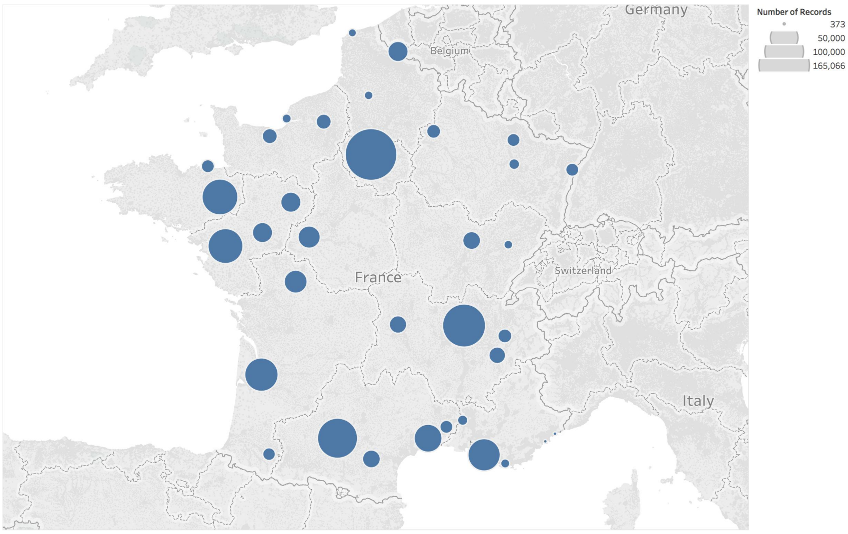


Figure 1: The map shows the volumes of records arriving to the city by the size of circle. Paris is the most popular destination while Lyon and Toulouse follow.

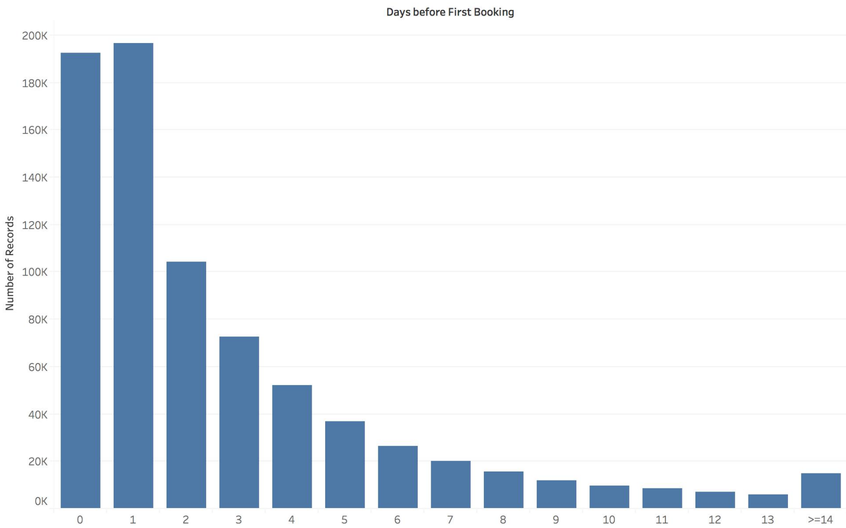


Figure 2: Most of the passengers booked their trips on the day of departure or one day before departure.

default price will be suggested to them.[6] There are also upper and lower bound price caps which prevent drivers from charging too high or too low prices. We collected the default price as well as the upper and lower price caps of each selected route by simulating driver's trip publication procedure using a driver's account (See Appendix D).

The final piece of the information is the strike rate computed and announced by SNCF. For most strike days, SNCF published a press document with overall strike rate. [7] We collected all the available strike rates and fill the missing values with the average of the strike date before and after. For non-strike dates, the strike rate was zero. The schedule of strike dates was well communicated before the strike started. We created an indicator of strike for trips departed on strike days.

5.1.3 Data cleaning and de-bias

At the end of the scraping process, we have an unbalanced panel data of trips departing from the end of March to the beginning of August 2018 which belong to the pre-defined 82 routes. The data is panel because a trip may have several observations up to 15 days before its departure date. The data is also imbalanced because a trip can no longer be traced once its departure date passes or that it be wiped out from the API. The fact of being able to trace back up to 15 days before departure allows us to observe the booking evolution, which is essential to demand curve estimation and welfare analysis. Changes in other important variables such as price level and total proposed seats can also be traced.[8] Of course, there could be no change in any variable at all during several days or for all observations of the trip.

For the convenience of analysis, we need to keep only one observation of each trip and create additional variables for the changes. In stead of keeping 15 observations of trip K that has 3 bookings happened in date A, B and C (we can know that by observing at which dates the seats left variable changed), it would be enough to only keep its last observation and indicate the total booked seats and the dates of booking. The same logic applied to changes in other variables. We firstly cleaned our dataset in this way to only keep one observation of each trip and add variable change indicators. This helped us significantly reduce the size of data while keep important information.

However, the data scraped from API is biased towards present final status of trips, even if the information scraped is correct. The main reason of bias is that the API does not keep trips forever. Several reasons can cause the trip to disappear or to be wiped our from API. If a trip is cancelled by the driver, it will no longer be found the next scraping day, even if the departure day has not come yet.[9] If the trip has been fully booked, there is also chance that the website stops showing the trip to potential passengers and that the trip is also wiped out from API. If the departure date of the trip has passed, the trip may also not be found on API the day after.[10]

This creates two biases. Firstly, we do not know if the trip disappears because of driver cancellation (no demand nor supply) or because of full booking (full demand and supply). Secondly, we are not sure if the booking number in the last observation of the trip is the true last status. That is to say, if there is no change in booking between the last scraping before departure and the wipe off of the trip from API. We scrap trips once a day which may allow non-captured changes

[6] On the website, if a driver wants to raise price up from the default level, she will be warned that most of the trips are cheaper and that there are more chances to stick to the default price level to maximize the chance of being booked. If she still want to increase the price level, the color of the price will switch from green to orange and eventually to red. Setting price lower than default level will not trigger warning message or change of color. On the mobile app, the color of the price never changes, but drivers do receive the warning and do need to double confirm before being able to modify the price level.

[7] Information is extracted here : https://www.sncf.com/fr/groupe-sncf/newsroom/communique-de-presse

[8] In fact, before a booking happens, drivers are free to modify any information of the trip, including price level, available seat number, trip description, correspondence city, etc. Once a booking occurs, drivers can no longer modify price unless the booked seat is cancelled. See https://www.BlaBlaCar.fr/faq/question/comment-modifier-mon-annonce-avant-et-apres-une-reservation for more information on trip modification.

[9] Drivers can cancel the trip at any time even if the trip has been booked. No penalty is applied except that drivers who often cancel at the last minute may not be able to publish new trips.

[10] Trip cancellation will lead to immediate wipe out from API, but fully booked or outdated trips could either be wiped out within one day, or keep staying on the API for a few days. We have not figured out how exactly the length kept is decided, which also adds difficulty in backward inferring the reason of wipe out.

between two scrapings.[11] Even though we could not be sure of the final status of each trip, we could nevertheless apply some rules to minimize the bias. Three scenarios may happen.

Scenario 1: At scraping date n+1, we still observe trips whose departure date is n. This is the ideal situation. We can take the booking data of the last scraping date as the final booking status of the trip.

Scenario 2: The last time that a trip is scraped is in the same day as its departure date. This situation is less ideal but still gives us confidence on the data trustworthy of the last scraping. Since scraping happens at around 19 p.m. and lasts about one hour each day, trips departing before could all be seen finished when scraping happens. Thus, we can take the API data as true last booking status. For trips departing after 20 p.m. but somehow can no longer be found the day afterwards, we consider them be fully booked at the last minute and then be wiped out from the API. We then modify the demand amount in our dataset. [12]

Scenario 3: The last time that a trip is scraped is before its departure date. This is a most complex situation. Whichever rules applied, there must still be bias. We here apply a simple, straightforward but still efficient rule based on two rationales: Firstly, the earlier the trip disappears from API comparing to the departure date, the more likely that the driver cancels it. Secondly, the fewer seats are booked before the trip disappears from API, the more likely that the driver cancels it. It is especially likely the case when the trip disappears with no seat booked. Our rules of reestimating seats supplied and booked are as follows:

If a trip has 4 and more unbooked seats before being wiped out from API, even if it is only one day before departure, we assume that the driver has cancelled it. If a trip has 3 unbooked seats before being wiped out and only 3 seats are supplied, we also assume that it has been cancelled. However, if total supplied seats are greater than 3, which means that the trip has had booking before being wiped out, we assume that it has been fully booked. If a trip has only 1 or 2 unbooked seats before being wiped out, we consider it as fully booked because it is relatively easy to have one or two people booking at the last minute.

5.2 Summary Statistics

Our dataset contains (almost) all trips information of 82 routes from the 1st of April 2018 to the 31st of July 2018. As illustrated by Figure 1, we cover all big cities in France and some intermediate and small cities. Paris is most popular destination, followed by Lyon and Toulouse. Figure 3 shows the total amount of trip records per route.[13] The busiest route is Nantes-Rennes, reaching almost 32,000 in the four-month period. Figure 4 shows the trip records of Paris-Lyon over the period and strike days are highlighted in yellow. We find both weekday effects and month effects. Roughly speaking, we find the amount fluctuating a lot in the beginning period of the strike (early April to mid May), and becomes stable in the second half of the period (mid May to late June). To better compare the difference between strike and non-strike days, we compute the average number of trips, as shown in Table 1. On average, 8,711 and 8,310 trips were offered on BlaBlaCar on a strike day and a non-strike day respectively, an increase of 4.8%.

Table 1: Average Number of Trips per Day

	Mean	SD	Min	Max
Non-Strike	8310	3112	4745	16728
Strike	8711	3536	4935	17804

[11] There may even be trips that are published, fully booked and then wiped out between two scraping dates, that we are even not able to scrap them once. They are considered as have never existed which will make our estimation of supply and demand downward biased. However, we ignore this bias in this paper because we have no way to verify its scope nor to correct it.

[12] There is also possibility that the driver cancelled the trip in the last minute, but since it should be rare and not encouraged by the platform, we ignore this possibility here.

[13] Only half of the pairs, i.e. 41 routes, are shown.

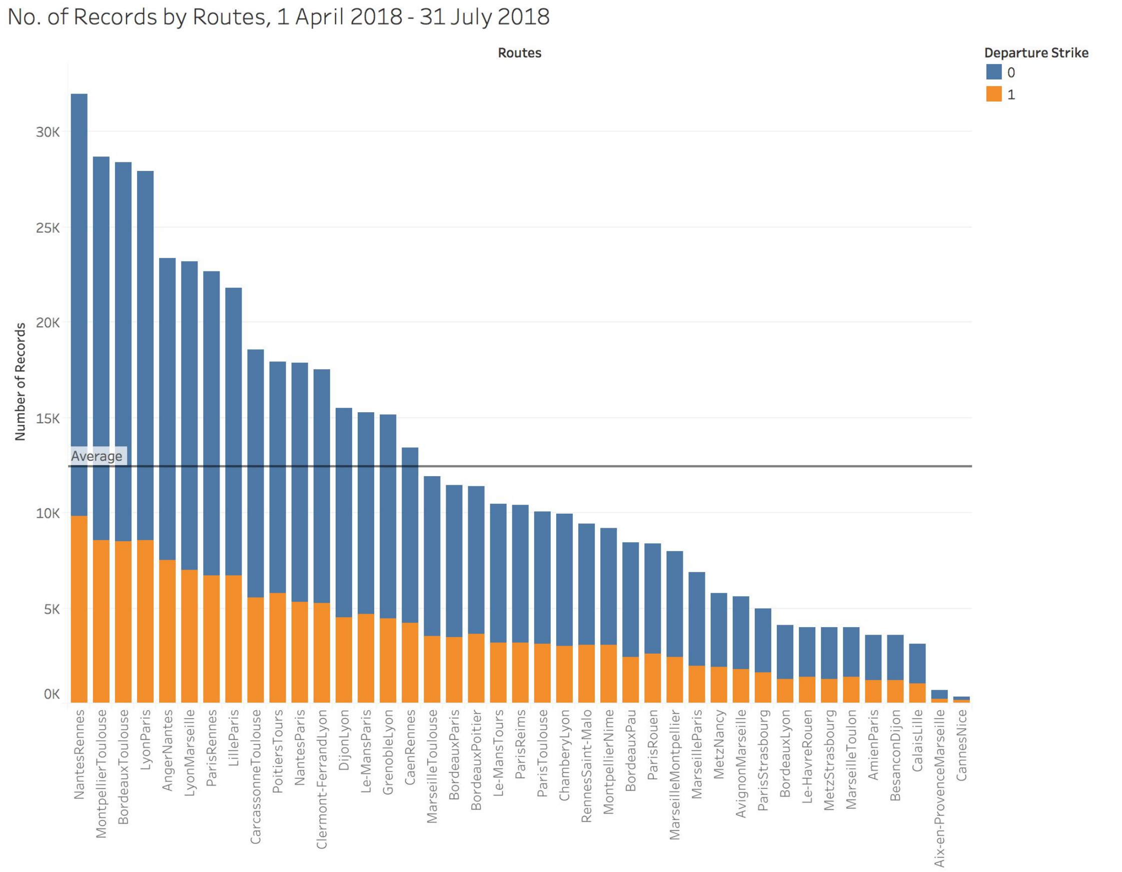


Figure 3: Routes are ranking according to the total trip records. Nantes-Rennes ranks the first, followed by Montpelier-Toulouse and Bordeaux-Toulouse. The least busiest routes are Cannes-Nice. The average amount of trip records is 12,465.

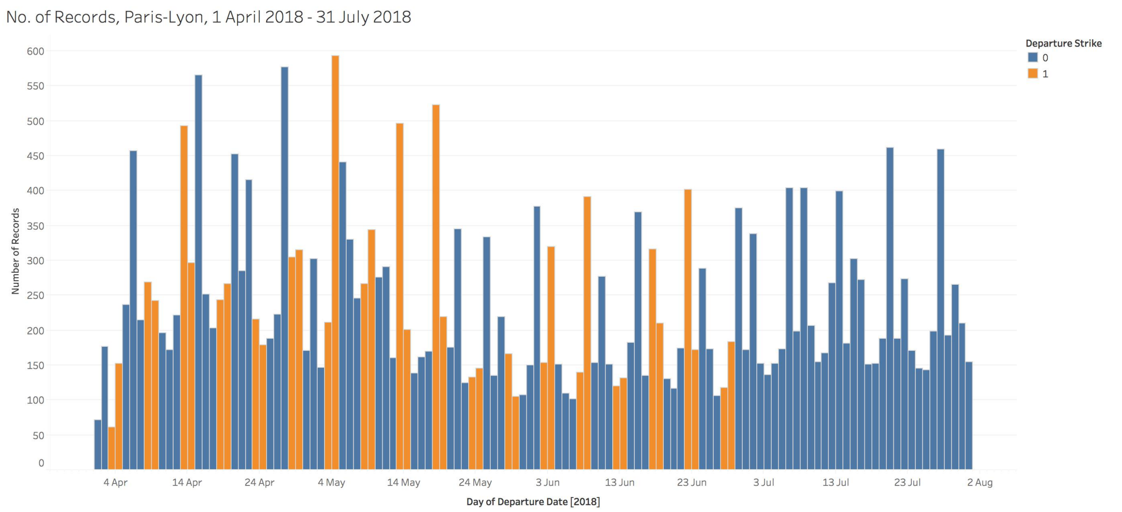


Figure 4: The numbers of trip records each day from 1 April to 31 July are shown along with highight of strike days in yellow. The busiest day, 4 May, registered 593 trips.

6 Effects on Supply

One of the distinctive advantage of the dataset is the abundance of details. We observe the number of seats offered to the market by drivers along with prices. Each observation is thus a revelation of a driver's willingness to offer. Note that the pair of price and quantity is not necessary a transaction while most of the conventional datasets contain only transactions. But we do not observe the willingness to pay of a driver over a range of prices. To complete the drawing of the supply curve, we make one essential assumption. Drivers are assumed to be willing to offer the same amount of seats at any price higher than the observed price but not lower. An individual driver during a short period of time can only offer a fixed amount of seats and thus an individual supply curve is perfectly inelastic over a large range of prices. If we consider that drivers have no outside option (no opportunity cost) and they have to drive in any case, an individual supply curve is a vertical line beginning from the price set by the driver, as shown in Figure 5.

The market aggregate supply is the horizontal summation of individual supply curves, as illustrated in a simplified way in Figure 6. We add up all the individual supply curves for each route on a daily basis. Although the services or the seats may not be directly competing with each other because passengers may prefer departing at certain time or taking a ride by a driver of certain gender, we ignore these minor details and define the market supply curve within the 24 hours of a calendar day. An example is shown in Figure 7. We observe that there is always a spike of supply at the reference price. Each route has a default price, which is recommended by BlaBlaCar, but drivers can always adjust the price within a range of prices. For example, the reference price for Paris-Lyon is 30 €. We can reasonably assume that some drivers quoting 30 € in fact would have accepted a price less than 30 €. That implies that the increase in quantity supplied below 30 € would have been less sharp and the actual supply curve would lie to right of the supply curve we drawn for the section below the reference price.

Our next step is to study the impact of the SNCF strike on the supply environment. From what we observe, the relationship between price and quantity supplied is cubic. The cubic relationship remains in log-log space. As we model the supply and estimate an average estimate of the impact of strike, we include into the regression the cubic, the quadratic and also the log of price at level. Precisely, we estimate the following model:

$$ln(Qs_{it}) = \alpha + \beta_1 ln(P_{it}) + \beta_2 ln(P_{it})^2 + \beta_3 ln(P_{it})^3 + \beta_4 Strike_t + X\gamma + \epsilon_{it} \tag{1}$$

where Qs_{it} is the quantity supplied, P_{it} is the price of route i and day t, and X refers to the matrix of control variables, that includes route fixed effects, month fixed effects, and weekday fixed effects. The error term ϵ_{it} is assumed to be randomly distributed with mean zero. Results are shown in Table 2. Column (1) is the estimation result of Equation 1. The coefficient of the strike day indicator is positive and significant. On average, a strike day experienced 9.4% increase in supply. In Column (2), we interact price (level, quadratic and cubic) with routes. The magnitude falls to 8.6%. Column (3) replaces the strike day indicator by the overall strike rate published by SNCF, while keeping all interaction terms added in Column (2). A higher strike rate means that more staff went on strike during a given day. A increase of strike rate by 1 percentage point is correlated with 0.44% increase in supply. The average strike rate during the period is approximately 20%, which implies on an average strike day the impact is 8.9%.

There are at least two reasons behind the increase in supply. First, affected passengers of SNCF might dive their own cars and also offer their empty seats on BlaBlaCar. Second, some drivers wanted to take advantage of the strike to make some extra money. The impact is however relatively small compare to the impact on demand, as we shall see in next section.

Table 2: BlaBlaCar Supply during SNCF Strike

	(1)	(2)	(3)
Strike Day	0.0938***	0.0857***	
	(0.0107)	(0.0080)	
Strike Day (overall)			0.4425***
			(0.0399)
Route FE	Yes	Yes	Yes
Weekday FE	Yes	Yes	Yes
Month FE	Yes	Yes	Yes
N	73233	73233	73233
R^2	0.793	0.886	0.886
No. of Groups	82	82	82

Robust standard errors in parentheses

* $p < .1$, ** $p < .05$, *** $p < .01$

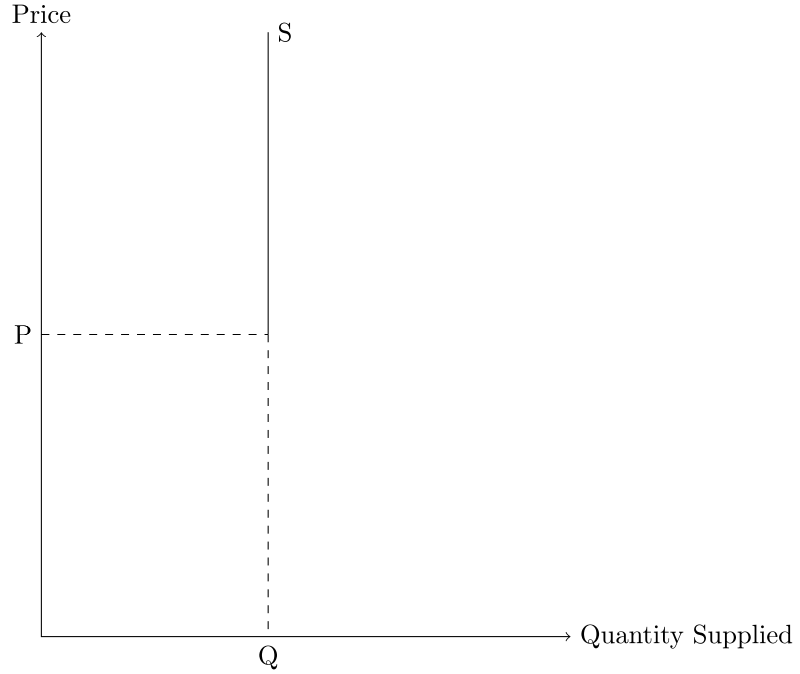


Figure 5: An individual supply curve is assumed to be a vertical line at Q beginning at the observed price explicitly chosen by the driver P.

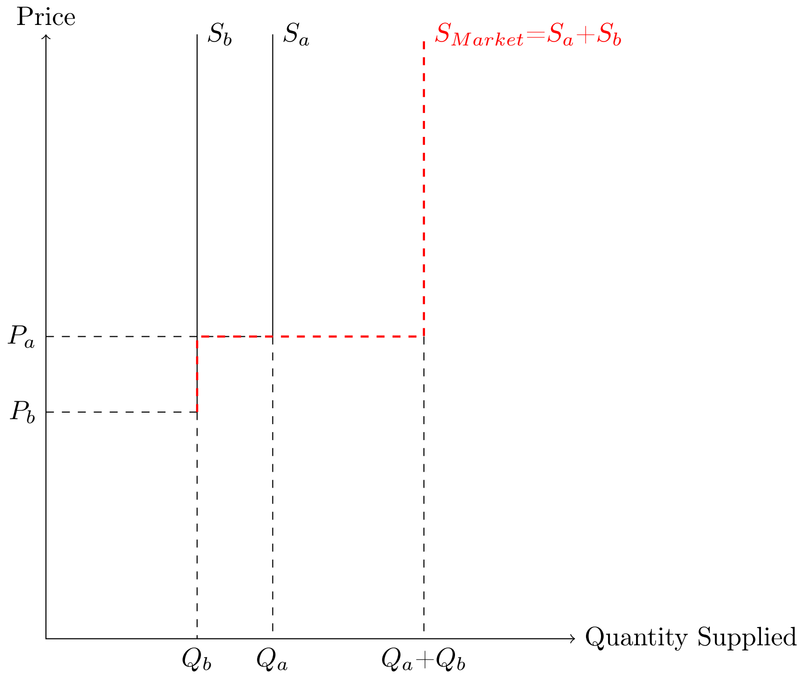


Figure 6: This graph illustrates how we horizontally sum up individual supply curves. Individual A(B) offers Q_a (Q_b) amount of seats. Between P_a and P_b, only individual B is willing offer and thus the market supply curve is the same as the individual supply curve of individual B. Beyond P_a, the market supply is Q_a+Q_b. Therefore, the market supply curve is represented by the dashed red line in this two-individual case.

7 Effects on Observed Demand

The construction of the market demand curve is less straightforward. We only observe transactions in the dataset but not the willingness to pay at all possible prices. For example, we cannot observe the willingness to pay when someone has not found a suitable ride. We first try to obtain the lower bound of the actual market demand by assuming that the maximum willingness to pay is equal to the price a passenger pay. The actual demand curve must lie on the right of the observed demand curve because passengers are very likely be willing to pay more than the price they pay. Moreover, we are unable to observe users' willingness to pay when they are unable to find a ride at certain prices. For simplicity, let us call the constructed curve below the observed demand curve.

Passengers have unit demand. Even if the price is zero, one does not need to transport two of oneself. By assuming that individuals would have booked the same trip at lower prices but not higher, an individual demand curve is thus a vertical line beginning at zero price and ending at the price the passenger paid. The aggregate observed market demand is the horizontal summation of all the individual demand curves.

We are interested in the impact of strike on the observed market demand. As the relationship between price and quantity demanded is also cubic, we estimate the following equation:

$$ln(Qd_{it}) = \alpha + \beta_1 ln(P_{it}) + \beta_2 ln(P_{it})^2 + \beta_3 ln(P_{it})^3 + \beta_4 Strike_t + X\gamma + \epsilon_{it} \quad (2)$$

Results are shown in Table 3. Column (1) reports the estimation of the baseline result. On average, observed market demand, or total payment, increased by 36.5% at each price level on a strike day. Column (2) reports the result of estimation that includes interaction terms between price (level, quadratic and cubic) and routes. The magnitude is almost the same, while R-squared improves. We then replace the strike indicator by the overall strike rate, as shown in Column (3). On an average day having 20% staff on strike, the impact on the observed market demand

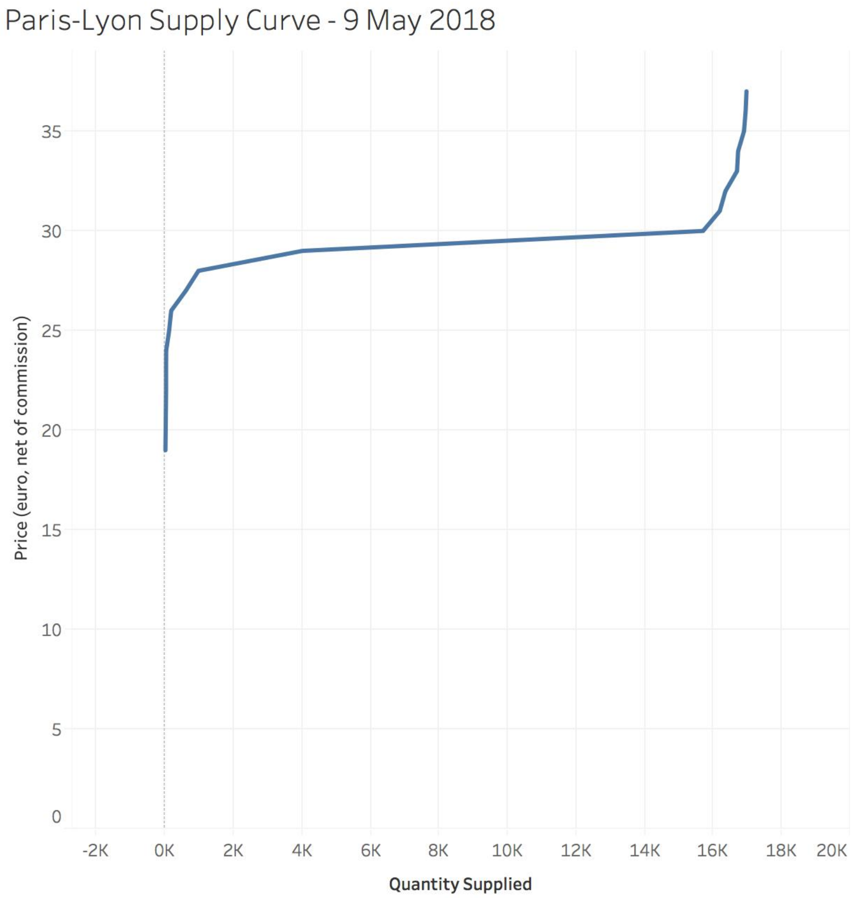


Figure 7: The market supply of the route Paris-Lyon on 9 May 2018 is depicted.

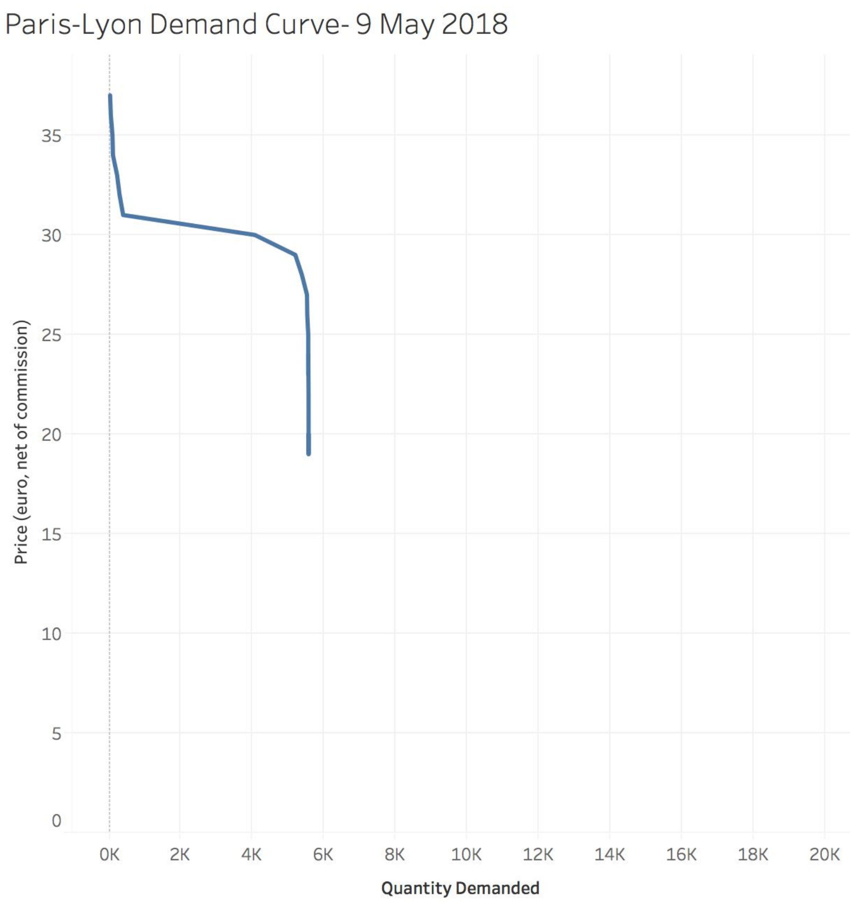


Figure 8: The observed market demand of the route Paris-Lyon on 9 May 2018 is depicted. As we do not observe any transactions at a price lower than 19 €, we complete the observed market demand curve by extending the curve downward vertically to zero. The actual demand curve should lie to the right of the constructed demand curve.

is roughly 37.3%. The increase in the observed demand due to the strike is substantial, while the supply is relatively inelastic.

Table 3: BlaBlaCar Demand during SNCF Strike

	(1)	(2)	(3)
Strike Day	0.3652***	0.3661***	
	(0.0110)	(0.0076)	
Strike Rate (overall)			1.8670***
			(0.0377)
Route FE	Yes	Yes	Yes
Weekday FE	Yes	Yes	Yes
Month FE	Yes	Yes	Yes
N	73249	73249	73249
R^2	0.735	0.876	0.877
No. of Groups	82	82	82

Robust Standard errors in parentheses

* $p < .1$, ** $p < .05$, *** $p < .01$

8 Welfare Analysis

8.1 Increase in Transaction Value

This section attempts to estimate the welfare impact due to BlaBlaCar on strike days. Perhaps the sentence is misleading. SNCF strike certainly led to some social welfare loss, while BlaBlaCar as a substitute salvaged some economic value. Nevertheless, we do not expect the net gain to be positive. As we are unable to obtain detailed information of train schedules of SNCF during the strike period, we are unable to measure the welfare loss due to the strike. However, as we have all transaction information of BlaBlaCar for the 82 routes, we can compare the payments between an average strike day and an average non-strike day and thus obtain an estimate of value transferred from passengers to drivers during the strike period, which is also a relevant indicator of how BlaBlaCar contributed to or alleviated the transportation chaos during the strike period.

We first compute the areas below each observed market demand curve of a route of a day, which are taken as the dependent variable to be explained by strike day indicator. The logic is illustrated by Figure 9. Imagine that $D_{non-strike}$ and D_{strike} are the observed market demand curves for a particular route on an average non-strike day and an average strike day respectively. As drivers can only set a price within a range set by BlaBlaCar and we construct the observed market demand curve by transaction information, we can only estimate the shape of the curve between P_{upper} and P_{lower} but not beyond them. The difference in the area below the curve, in this case a positive change, is the additional payment from passengers to drivers on the platform of BlaBlaCar on an average strike day.

Precisely, we estimate the following model:

$$ln(V_{it}) = a + b_1 Strike_t + Zc + e_{it} \tag{3}$$

where V_{it} is the area under the observed market demand curve of route i on day t, $Strike_t$ is the strike day indicator, matrix Z contains control variables including a linear time trend, route fixed effects, month fixed effects, and weekday fixed effects. The error term e_{it} is randomly distributed with zero mean. We are interested in the coefficient of $Strike_t$, which is interpreted as the impact of strike on the observed market demand on an average strike day. Results are shown in Table 5. Column (1) reports the baseline result where all routes share the same time trend. A strike day saw on average an increase of payment by 31.3% at each price level. Column (2) replaces the common time trend by route-specific trends but the estimate is almost identical. Column (3) replaces the strike day indicator by overall strike rate. On an average strike day with 20% staff on strike, the impact is 32%. The result is robust and significant. From April to July 2018, for the

82 routes we have selected, average total payments exchanged on the platform of BlaBlaCar on an average non-strike day is approximately 2.2 € million. An increase of 30% is equivalent to 0.70 € million. A summary of the estimates of total transaction value is shown in Table 4.

Table 4: Summary of Estimates of Total Transaction Value in millions euro (April-July)

	Total (April-July)	Average per day
Non-Strike Days	188.4	2.19
Strike Days	103.0	2.86

The press release by SNCF states that "At the end of June, SNCF Group's revenues stood at 16.6 € billion, down 3.3%, with TGV high-speed rail traffic down 3.8%. Without the strike, revenue would have risen by around 4%." In other words, the strike effect on revenue is 1.21 € billion for 37 days. On average, the loss is 3.27 € million per day. Let this amount be the economic loss due to the strike per day. BlaBlaCar recovered 21.4% of the loss, if we assume that the opportunity costs of drivers are zero.[14] Taking loss of revenue as the reference may be misleading as a strike does not simply imply tickets to be refunded or revenue unearned. SNCF spent extra money to provide assistance, chartering coach services, and so on. SNCF claims that roughly 21 € million was lost each strike day.

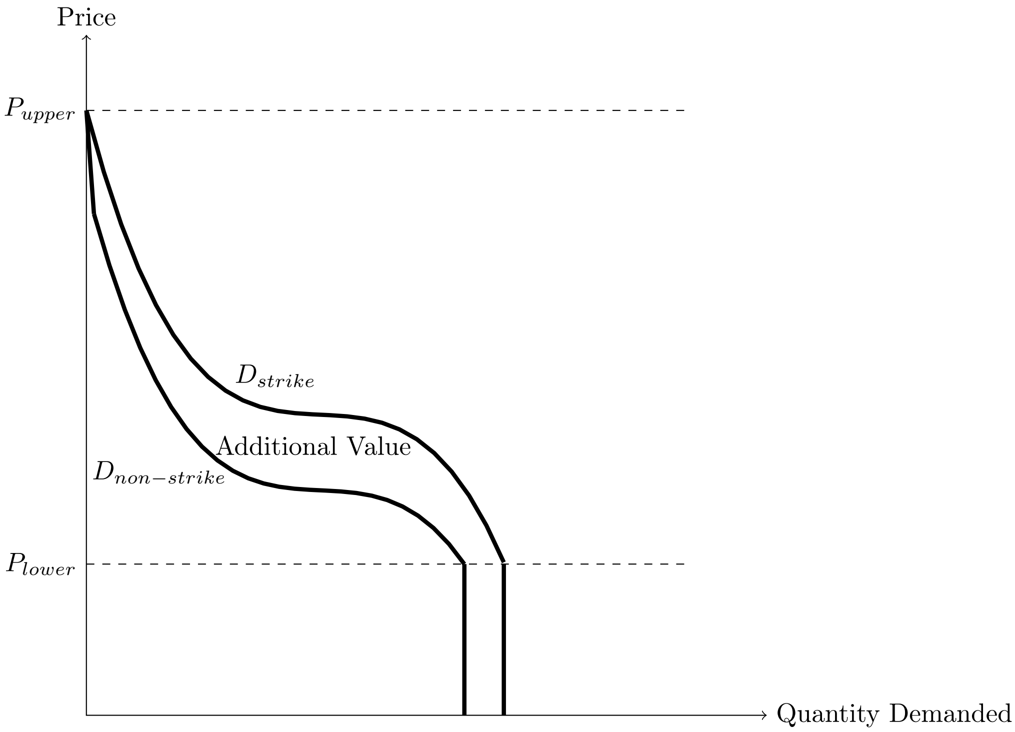


Figure 9: Observed market demand curves of an average strike day and an average non-strike day are depicted by D_{strike} and $D_{non-strike}$ respectively. The difference in the area below the observed market demand curve is the additional value generated by BlaBlaCar during an average strike day compared to an average non-strike day.

8.2 Effect on Consumer Surplus

In the analysis above, we assume that passengers' maximum willingness to pay coincides with the prices they pay. Using total payment as a measure of economic contribution may not be satisfactory because we have not considered costs on the supply side. Another widely accepted measure is

[14] https://www.sncf.com/sncv1/ressources/reports/pr_sncf_group_h1_2018_results_07.27.2018.pdf .Retrieved on 13 January 2019.

Table 5: Additional Payments Transacted on BlaBlaCar during SNCF Strike

	(1)	(2)	(3)
Strike Day	0.3128***	0.3122***	
	(0.0231)	(0.0230)	
Strike Rate (overall)			1.5995***
			(0.1202)
Linear Time Trend	Yes	Route-specific	Route-specific
Route FE	Yes	Yes	Yes
Weekday FE	Yes	Yes	Yes
Month FE	Yes	Yes	Yes
N	9966	9966	9966
R^2	0.400	0.426	0.429
No. of Routes	82	82	82

Robust standard errors in parentheses

* $p < .1$, ** $p < .05$, *** $p < .01$

consumer surplus, which is the distance between the maximum willingness to pay and the actual payment by consumers at each quantity. If we assume that all the payments simply cover the costs of the trips, the increase in consumer surplus is also the additional social welfare to the economy.

Standard approach to measure consumer surplus relies on equilibrium analysis. By assuming that all transactions observed are equilibrium intersections between demand and supply, researchers estimate price elasticity of demand using different identification strategies. Farronato and Fradkin (2018), using data on Airbnb, computed the average price of bookings and took it as the market clearing (equilibrium) price of each day to estimate price elasticity. The market demand curve is thus built upon average price of a day and total bookings of a day. However, we believe that with digital platform data we could avoid appealing to aggregation of information. Structural models rely on equilibrium, which is often assumed to coincide with an observed price-quantity pair. It is reasonable as we may only observe at a time the price set by the seller and the total transaction amount of the product at the price in most of the brick-and-mortar markets, and, therefore, the resulting demand curve is a collection of points of different observation periods. The construction of the demand curve is thus not immune to endogeneity bias as equilibrium price and quantity transacted are the result of interactions between demand and supply. Traditional econometrics of demand estimation has been built around this implicit assumption.

Cohen et al. (2016) stepped back from traditional approach and relied on discontinuity to produce some randomness around a price. As they observed acceptance and rejection of price offers by Uber users, they could estimate the price elasticity of demand around those discontinuities and built the market demand curve. Their result is immune to many challenges since they also observe rejections of offers. Although BlaBlaCar does not have discontinuities due to price surge as Uber, our approach, to a certain extent, is similar to the one by Cohen et al. (2016). We do not appeal to equilibrium analysis, and construct quantities demanded and supplied at each available price by exploiting the richness of the platform data. We then establish the daily observed market demand curve and the market supply curve. To arrive to an estimate of the consumer surplus, the basic assumption made is that the observed demand curve gives us information to obtain the true price elasticity, together with a conservative estimate of the intercept of the market demand curve.

As we observe all the payments but not the maximum willingness to pay, we need to make assumptions before reaching any conclusion. Assume a constant price elasticity demand curve, we write the demand function as follows:

$$Q_{itp} = K_{it} p^{\eta} \tag{4}$$

where Q_{itp} is the quantity demanded at price level p of route i on day t, K_{it} capture all other variables except Q_{itp} and η is the price elasticity of demand. If K_{it} is to be shown on a usual price-quantity dimension, it is the x-intercept of the demand curve. Although the constant-elasticity assumption tends to oversimplify the real world, the advantage is that once we obtain a reliable estimate of the price elasticity of demand we are able to draw the whole demand curve. The consumer surplus is thus the difference between the area below the observed demand curve we constructed in Section 7, which is exactly the total payment of all transactions, and the area below the estimated demand curve of Equation 4.

8.2.1 Optimistic Estimation

As observed transactions are affected by the supply condition that is manipulated by the recommended price, we observe spikes at the recommended price for both the demand and supply curves. If we estimate the price elasticity around the reference price using the actual transaction data, we will find a drastic change due to the spike, contrasting to our assumption. At prices slightly above and below the reference price, for example, two euros above and two euros below, there are reasonable amounts of offers and transactions available where the estimate of the elasticity is more reliable. We thus exclude all observations at the reference price and those too far from the reference price. Precisely, we only include observations within the ranges $[P_{ref}+2, P_{ref})$ and $(P_{ref}, P_{ref}-2]$, and estimate the log of Equation 4 to obtain a single estimate of the elasticity for all routes while controlling for weekday, month and route fixed effects, along with a linear time trend and an indicator of the section of the demand curve (above or below the reference price). This method will be referred as Method A below. Results are shown in Table 6.

Table 6: Elasticity Estimation

	(1)	(2)	(3)
η	-2.8428***	-2.7735***	-2.7740***
	(0.0857)	(0.0853)	(0.0854)
Strike Day	0.3520***	0.3520***	
	(0.0132)	(0.0131)	
Strike Rate (overall)			1.7689***
			(0.0642)
Linear Time Trend	Yes	Route-specific	Route-specific
Route FE	Yes	Yes	Yes
Weekday FE	Yes	Yes	Yes
Month FE	Yes	Yes	Yes
N	32870	32870	32870
R^2	0.797	0.801	0.802
No. of Routes	82	82	82

Standard errors in parentheses

* $p < .1$, ** $p < .05$, *** $p < .01$

The first column models a common linear time trend for all routes. The estimate of the elasticity η is -2.84, meaning that one percent increase in price leads to a fall of quantity demanded by 2.84%. Strike day indicator is associated with a positive coefficient, as expected, meaning that the market demand of an average strike day lies to the right of the market demand of an average non-strike day. Column (2) instead assumes route-specific time trends, but the estimates are similar. Column (3) replace the indicator by the overall strike rate. Again the estimates are robust. To compute the consumer surplus, we take η as -2.7735.

Regression of Equation 4 will give us an estimate of the elasticity, but we need to transform the equation for us to easily compute the consumer surplus as the following:

$$p = \left(\frac{Q_{it}}{K_{it}}\right)^{\frac{1}{\eta}} \tag{5}$$

Consumer surplus is computed as follows:

$$C_{it} = \int_0^{\overline{Q}_{it}} (\frac{Q_{it}}{K_{it}})^{\frac{1}{\eta}} dQ - V_{it} \tag{6}$$

where $\overline{Q}_{itp}$ is the max quantity demanded at the lower bound price, and V_{it} is the value below the observed demand curve we constructed in Section 4. Equation 7 can be transformed as follows:

$$C_{it} = \frac{1}{K_{it}^{\frac{1}{\eta}}} \frac{1}{\frac{1}{\eta}+1} \overline{Q}_{it}^{\frac{1}{\eta}+1} - V_{it} \tag{7}$$

By replacing K_{it} and η by therir corresponding estimates $\hat{K}_{it}$ and $\hat{\eta}$, we obtain an estimate of the consumer surplus for each route on each day. A graphical illustration is given in Figure 10. What left to determine is the intercept term $1/K_{it}^{\frac{1}{\eta}}$, i.e. the price at which the first quantity demanded would have been conceived. We proceed as the following. We take the predictions of K_{it} obtained from the regression of the log of Equation 4 as a baseline $\hat{K}_{it}$. Next, we multiply the baseline $1/\hat{K}_{it}^{\frac{1}{\eta}}$ by a factor to force the consumer surplus of any route on any day at least zero because negative consumer surplus is inconceivable.[15] The factor is found to be 3.225. We then add up all estimated consumer surplus and summarize in Table 7. The difference in the estimated consumer surplus between strike and non-strike days can be interpreted as the additional welfare gain on the consumer side due to BlaBlaCar. Table 8 reports the estimations of the impact of strike on consumer surplus. Specifications remain the same as those in Table 5. We find that on average a day of strike led to an increase in consumer surplus by 35%, which is robust across three specifications. The result again suggests that there were substantial amount of substitutions happened during the strike period. One could interpret the result as BlaBlaCar salvaged on average 2 million per strike day.[16] To have a perspective of the size of the effect, we compared the estimate to the loss stated by SNCF, which is 21 million per strike day, and it is roughly 9.78% of the loss.

[15] As an indicator of whether the price is above or below the reference price is included in the regression to refrain the slope from being influenced by the spike at the reference price, we actually have two $\hat{K_{it}}$ for each route-day pair. In the following analysis we take the $\hat{K_{it}}$ associated with the prices below the demand curve.

[16] 5.87 X 0.35 = 2.0545

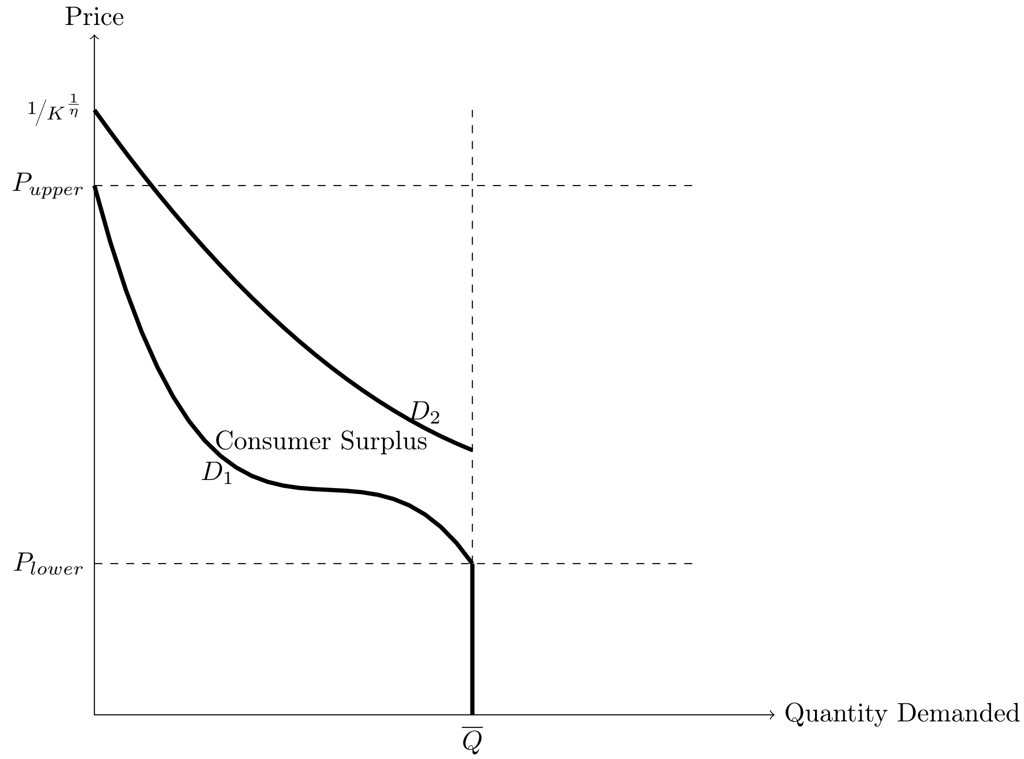


Figure 10: This graph illustrates the computation of the consumer surplus. The area under the observed market demand curve (D_1) is exactly the total payment of all bookings. By gathering observations around the reference price while excluding the reference price, we estimate the elasticity, which is then projected to a hypothetical market demand curve (D_2) given an approximation of K. The area between two curves is the consumer surplus.

Table 7: Summary of Estimates of Consumer Surplus in millions euro (April-July)

	Total (April-July)	Average per day	CS to Transactions
Non-Strike Days	504.8	5.87	3.46
Strike Days	282.8	7.86	3.45

Table 8: Additional Consumer Surplus due to BlaBlaCar during SNCF Strike

	(1)	(2)	(3)
Strike Day	0.3467***	0.3462***	
	(0.0224)	(0.0223)	
Strike Rate (overall)			1.6354***
			(0.1258)
Linear Time Trend	Yes	Route-specific	Route-specific
Route FE	Yes	Yes	Yes
Weekday FE	Yes	Yes	Yes
Month FE	Yes	Yes	Yes
N	9954	9954	9954
R^2	0.511	0.556	0.556
No. of Routes	82	82	82

Robust standard errors in parentheses

* $p < .1$, ** $p < .05$, *** $p < .01$

We also compute the ratio of consumer surplus to total transaction value, shown in Table 7, the ratio is roughly 3.45-3.46 on an average day, meaning that on average a euro spent on BlaBlaCar generates 3.45-3.46 € economic value.

8.2.2 Conservative Estimation

The optimitis estimation requires that the sections of the observed market demand around the reference price while excluding the reference price are representative. But it may overestimate the price elasticity of demand. The sections of the observed market demand curve are very likely to be less elastic than the true ones because the supply condition, which is heavily biased towards the reference price due to the manipulation of BlaBlaCar, may suppress the quantities demanded that we can observe. We amend this problem as the following. First, we estimate one single estimate of the elasticity by linking observations of the two sections while still keeping the reference price out, referred as Method B. It is expected to produce a more elastic estimate, as Figure 11 shows. Second, we include also the information of supply in the regression while utilizing also observations at the reference price. Including the quantity supplied helps alleviate the impact of the manipulation at each price level. This method will be referred as Method C. For example, the spike of the supply captures the spike of the demand and we could be more confident to believe that the price elasticity of the actual market demand curve is a constant, given that the spike is well controlled. Third, we amend Method C by including an interaction term between supply and an indicator of observations equal or below the reference price, referred as Method D. Finally, we amend Method D by removing all observations at the reference price, referred as Method E. Results are shown in Table 9. Column (1) is a reproduction of the specification while excluding the section dummy. As expected, the estimate of elasticity is larger, implying a less step market demand curve and a lower consumer surplus. Column (2) includes the log of quantity supplied at the corresponding price level while including also the observations at the reference prices. The estimate is even larger and the coefficient of supply is highly significant. We believe that the coefficient is too large as the spike of demand seems not well modeled by an obviously wrong linear relationship between the log of supply and the log of demand. Besides, the impact of the supply on the demand depends on whether the corresponding price level is above or below the reference price. On top of the previous column, Column (3) includes an interaction term between the indicator of section below the reference price and the log of supply. The interaction term is highly significant, reflecting the drag by the reference price at the section of the demand at prices equal to or below the reference price. Column (4) removes all observations at the reference prices and re-estimate the model of Column (3) because the existence of the observations at the spike is likely to produce a very flat market demand curve despite the control of the supply condition. The estimate of elasticity seems to fall within a more reasonable range. The estimate is robust, ranging from -6.76 to -6.69, subject to changes of the specification of the model including replacing the common time trend by route-specific time trends and replacing the strike indicator by strike rate, though the details are not shown in the paper.

Table 9: Price Elasticity of Demand by Conservative Approaches

	(1) Method B	(2) Method C	(3) Method D	(4) Method E
η	-7.6502***	-11.8650***	-7.6464***	-6.7607***
	(0.0562)	(0.1027)	(0.1076)	(0.1198)
Strike Day	0.3527***	0.2852***	0.3062***	0.3140***
	(0.0148)	(0.0118)	(0.0100)	(0.0116)
ln Supply		0.4203***	0.1854***	0.2380***
		(0.0057)	(0.0077)	(0.0098)
≤Reference			-0.0098	0.6028***
			(0.0473)	(0.0575)
≤Reference X ln Supply			0.2118***	0.2007***
			(0.0060)	(0.0078)
N	32870	42263	42263	32870
R^2	0.743	0.781	0.843	0.841
No. of Routes	82	82	82	82

Robust standard errors in parentheses

* $p < .1$, ** $p < .05$, *** $p < .01$

Table 10: Summary of Estimates of Consumer Surplus in millions euro (April-July) by Conservative Approaches

	Total (April-July)	Average per day	CS to Transactions
Non-Strike Days	125.3	1.46	0.636
Strike Days	70.1	1.95	0.641

Table 11: Additional Consumer Surplus due to Blablacar during SNCF Strike by Conservative Approaches

	(1)	(2)	(3)
Strike Day	0.3313***	0.3307***	
	(0.0190)	(0.0189)	
Strike Rate (overall)			1.6153***
			(0.1012)
Linear Time Trend	Yes	Route-specific	Route-specific
Route FE	Yes	Yes	Yes
Weekday FE	Yes	Yes	Yes
Month FE	Yes	Yes	Yes
N	9966	9966	9966
R^2	0.486	0.508	0.510
No. of Routes	82	82	82

Robust standard errors in parentheses

* $p < .1$, ** $p < .05$, *** $p < .01$

9 Discussion

9.1 On Consumer Surplus Estimation

Readers may challenge that the proposed method to calculate consumer surplus is questionable since the observed demand curve, even if the manipulation by BlaBlaCar has been corrected by the information on the supply condition, may not truly reflect the price elasticity of the true market demand curve. We acknowledge some common criticisms, such as the simplistic form of demand curve, but argue that our proposed method at the very least gives a correct ranking of price elasticity, and thus consumer surplus. Figure 11 illustrates two hypothetical demand curves in log-log space obtained from estimating Equation 4, where the slope is exactly the price elasticity. On the left hand side, the two curves have not been corrected by the supply condition, while they are corrected on the right hand side. We first limit to the case that the two demand curves have the same prediction of K.

If we exclude the information at the reference price and estimate the slope of the two sections separately (Method A), we obtain the demand curves on the left hand side. The ranking of slope (price elasticity) is correct in a sense that consumer surplus is larger under D_2 when we observe more passengers are willing to pay a higher price for the same trip. If we include also the information at the reference price while not considering the supply condition (Method B), we obtain a much flatter slope that considers also the spike in quantity demanded. Still, the ranking is correct. When the supply condition is taken into account in the regression (Method C, D and E), the two sections are ideally matched together as shown by the right-hand side. Again, the estimate of the elasticity gives us a correct ranking of the consumer surplus of the two hypothetical demand situation because passengers are willing to pay more for the same unit of trip.

Another extreme case is that the two demand curves are sharing the same slope but not the same x-intercept (not graphically shown here). The ranking of the consumer surplus is also correct because we keep the ratios between intercepts of the observed demand curves constant while projecting the intercepts to obtain the estimated true intercepts.

Cases in between (when two demand curves cross) are a bit more complicated and there could be cases that we are unsure that which situation gives a higher consumer surplus. But the demand curve can always be divided into sections and each section falls closer to the illustrated extreme cases. To conclude, as no one really observes the maximum willingness to pay of consumers, all estimation techniques are only approximations. Our proposed methodology, which relies on actual data information and simple assumptions while avoiding guesses of parameters and calibration, is arguably more transparent, reliable and produces as good estimates as other techniques.

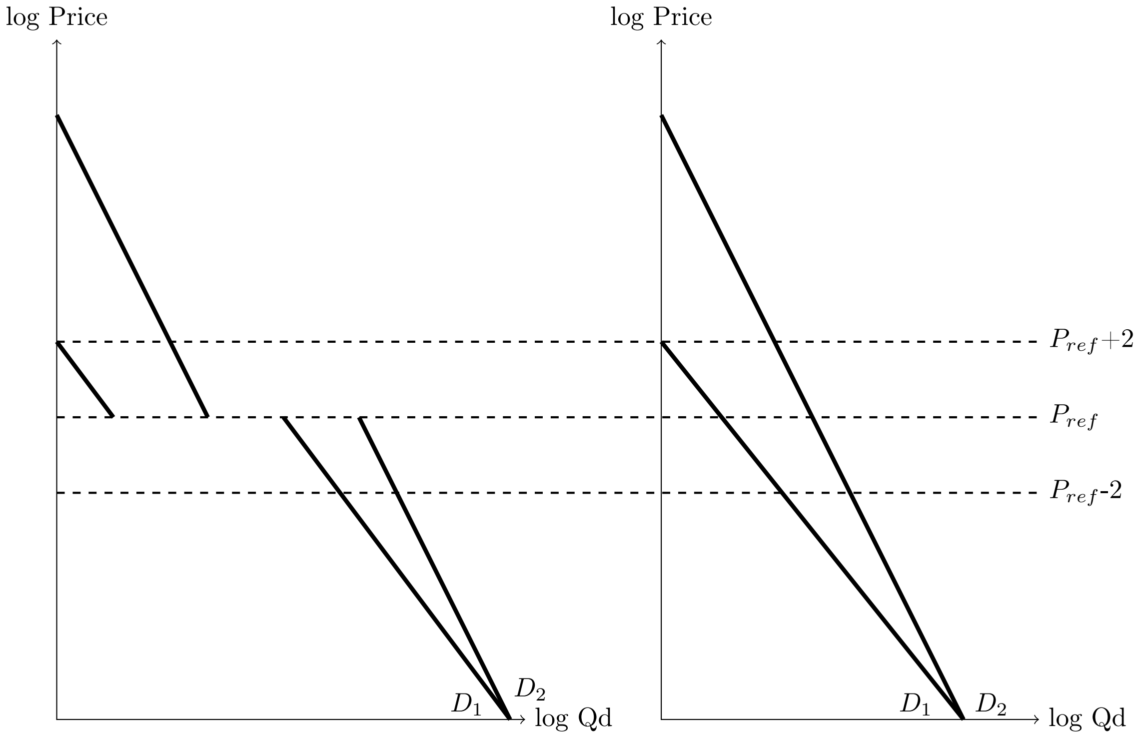


Figure 11: On the left hand side, the two curves have not been corrected by the supply condition, while they are corrected on the right hand side.

9.2 On Welfare Estimation

The data presented in the previous sections compare the scale of BlaBlaCar consumer surplus (passenger surplus) gain during strike with SNCF loss during strike. This demonstration, however, should not be naively interpreted as the percentage of SNCF loss recovered by BlaBlaCar, for two reasons. Firstly, not all SNCF loss is related to welfare loss of its users. Loss includes not is not limited to sales loss, refund to passengers as well as extra costs of customer assistance. Even for the sales loss and passenger refund parts which count for 160 million euros altogether, it is not equal to the welfare loss because the welfare loss of SNCF passengers during the strike is not included. Secondly, the welfare gain from using BlaBlaCar during strike cannot be reflected to SNCF welfare loss either. Passengers who use BlaBlaCar with or without strike benefit from a welfare gain as well. Some SNCF passengers became BlaBlaCar drivers during strike, whose welfare gain falls into the producer (driver) surplus part that we have not analyzed in this paper.

Another possible extension of the paper is on the channel of calculating welfare. For the moment, we have only estimated material welfare impact using price. Time is another dimension which may heavily influence welfare differently across route types. For routes between major cities where high speed railway (TGV) system is well developed, taking SNCF could save hours comparing to taking BlaBlaCar. For routes between two secondary cities close by, often times only normal speed trains (TER) are available and the time costs of the two modes are similar.

10 Conclusion

This work attempts to understand any effect of the SNCF strike from April to June 2018 on ridesharing behaviours. Using the information available on BlaBlaCar official API, we construct the market demand and supply of BlaBlaCar and find that the impact on demand was much larger than on supply. Our estimate shows that demand on average increased by 37% on a strike day, while supply increased by only 9%. We also provide a conservative and an optimistic estimate of the additional consumer surplus generated by BlaBlaCar as a platform between drivers and passengers on an average strike day, which reached 0.48 - 2 € million per strike day, or 2.30 - 9.78% of the loss of SNCF per strike day. Our results show substantial substitutions between SNCF and ridesharing on BlaBlaCar, providing evidence that the digital platform economy has significantly transformed our economy.

References

Chen, M. K. and Sheldon, M. (2015), Dynamic Pricing in a Labor Market: Surge Pricing and the Flexible Work on the Uber Platform.

Chevalier, J. A., Chen, M. K., Oehlsen, E. and Rossi, P. E. (2018), The Value of Flexible Work: Evidence from Uber Drivers.

Cohen, P., Hahn, R., Hall, J., Levitt, S. and Metcalfe, R. (2016), Using Big Data to Estimate Consumer Surplus: The Case of Uber.

Cook, C., Diamond, R., Hall, J., List, J. A. and Oyer, P. (2018), The Gender Earnings Gap in the Gig Economy: Evidence from over a Million Rideshare Drivers.

Crozet, Y. and Guihéry, L. (2018), 'Deregulation of Long Distance Coach Services in France', *Research in Transportation Economics* **69**, 284–289.

Exel, N. J. A. V. and Rietveld, P. (2001), 'Public Transport Strikes and Traveller Behaviour', *Transport Policy* **8**, 237–246.

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* .

Farronato, C. and Fradkin, A. (2018), The welfare effects of peer entry in the accommodation market: The case of airbnb, Technical report, National Bureau of Economic Research.

Fearnley, N., Currie, G., Flügel, S., Gregersen, F. A., Killi, M., Toner, J. and Wardman, M. (2018), 'Competition and Substitution Between Public Transport Modes', *Research in Transportation Economics* **69**, 51–58.

Ge, Y., Knittel, C. R., Mackenzie, D. and Zoepf, S. (2016), Racial and Gender Discrimination in Transportation Network Companies.

Hall, J. D., Palsson, C. and Price, J. (2018), 'Is Uber a Substitute or Complement for Public Transit ?', *Journal of Urban Economics* **108**, 36–50.

Kim, K., Baek, C. and Lee, J. D. (2018), 'Creative Destruction of the Sharing Economy in Action: The Case of Uber', *Transportation Research Part A: Policy and Practice* **110**, 118–127.

Lam, C. T. and Liu, M. (2017), Demand and Consumer Surplus in the On-demand Economy: the Case of Ride Sharing.

Larcom, S., Rauch, F. and Willems, T. (2017), 'The Benefits of Forced Experimentation: Striking Evidence from the London Underground Network', *The Quarterly Journal of Economics* **132**(4), 2019–2055.

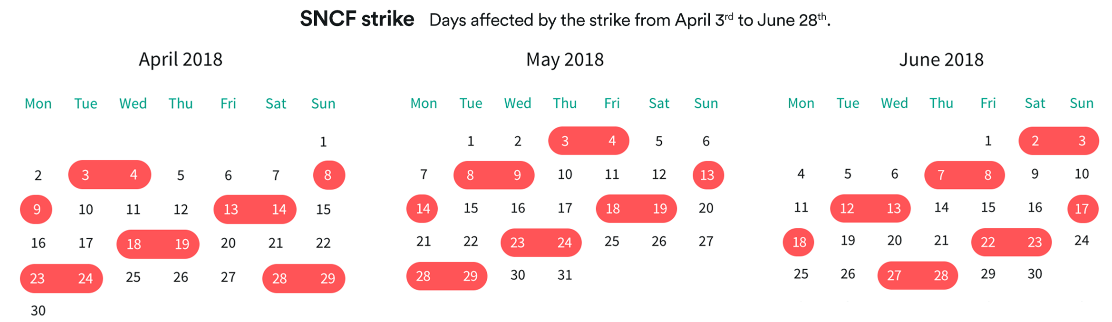
Schwieterman, J. and Smith, C. S. (2018), 'Sharing the ride: A paired-trip analysis of UberPool and Chicago Transit Authority services in Chicago, Illinois', *Research in Transportation Economics* **71**, 9–16.

Shaheen, S., Stocker, A. and Mundler, M. (2017), Online and App-Based Carpooling in France: Analyzing Users and Practices—A Study of Blablacar, *in* 'Disrupting Mobility', Springer International Publishing, pp. 181–196.

Zhu, S. and Levinson, D. M. (2012), Disruptions to Transportation Networks : A Review, *in* 'Network reliability in practice', Springer, New York, pp. 5–20.

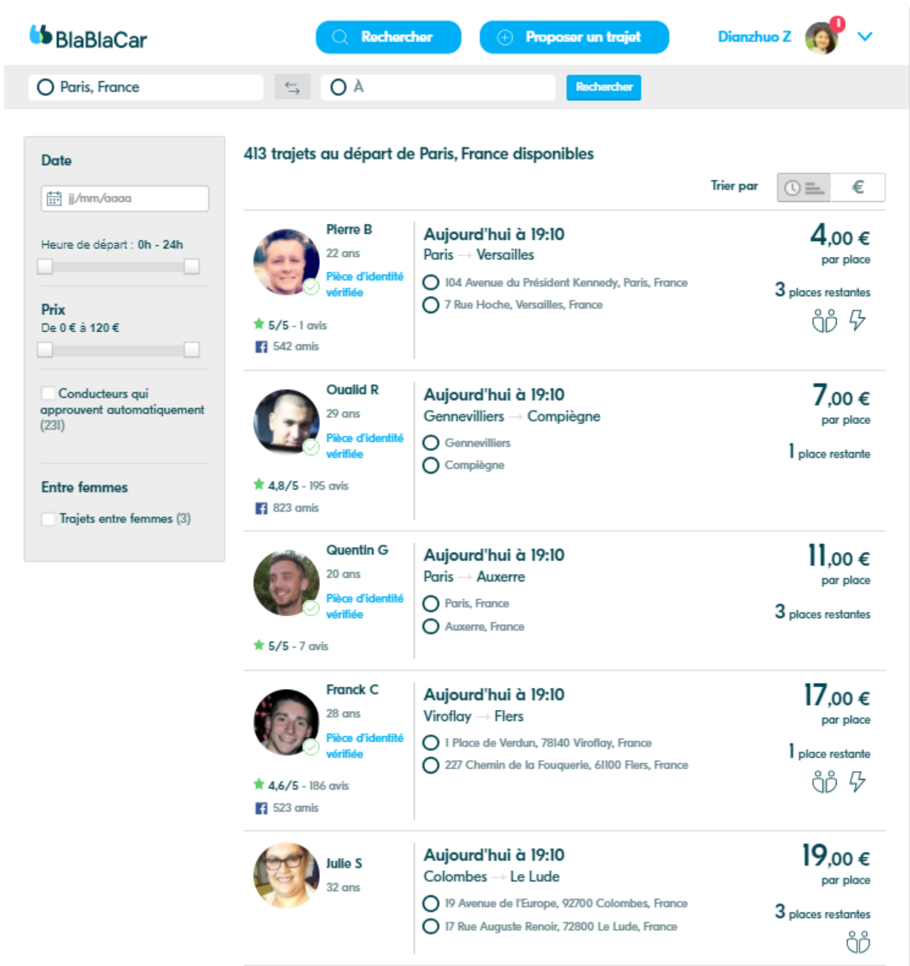
Appendices

A SNCF strike calendar



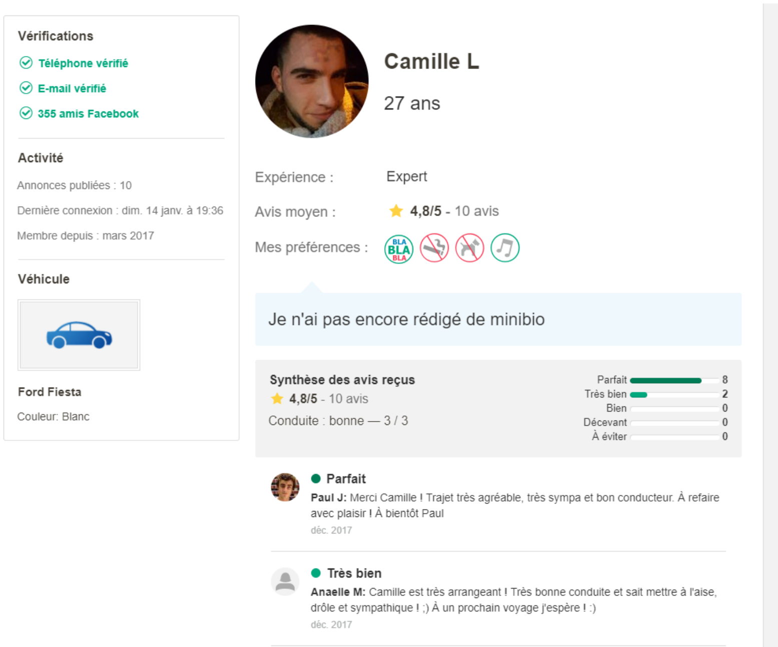
Strike dates are marked in red. Retrieved from https://faq.trainline.eu/article/674-sncf-french-rail-strikes-2018 in 14th January 2019.

B Schema of trip search of BlaBlaCar



Demonstration of trip searching result

Demonstration of trip information page and driver information page (below). Retrieved in January and Febuary 2018. All information were publically available.



Demonstration of driver information page. All demonstrations are retrieved in January and Febuary 2018. All information were publically available.

C Commission level of BlaBlaCar

Table 12: List of commission level charged by BlaBlaCar

Price charged by driver	Commission level
1-6	1.0
7-8	1.5
9-11	2.0
12-13	2.5
14-16	3.0
17-18	3.5
19-21	4.0
22-23	4.5
24-26	5.0
27-28	5.5
29-32	6.0
33-35	6.5
36-37	7.0
38-40	7.5
41-42	8.0
43-46	8.5
47-50	9.0
>51	18% of price charged by driver, rounded at 0.5

Priced are in euro. Commission level is not changing according to routes, hours or days. Information collected from https://blog.BlaBlaCar.fr/blablalife/lp/nouvelle-grille-de-frais-de-reservation

D Selected routes

Table 13: List of routes scrapped (one way) and reference information

Route	Upper price cap	Lower price cap	Default price	Distance
Major to major cities				
Paris Lyon	37	13	29	469
Paris Marseille	62	22	49	775
Paris Bordeaux	47	17	37	585
Paris Toulouse	54	20	43	679
Paris Strasbourg	39	14	31	492
Paris Nantes	30	11	24	385
Lyon Marseille	25	9	20	314
Lyon Bordeaux	45	16	36	556
Marseille Toulouse	33	12	26	403
Marseille Montpellier	14	5	11	170
Bordeaux Toulouse	19	7	15	246
Toulouse Montpellier	19	7	15	243
Paris Rennes	28	10	22	354
Major to secondary cities				
Paris Lille	18	6	14	219
Paris Amien	12	4	9	144
Paris Reims	11	4	9	144
Paris Rouen	11	4	9	136
Paris Le-Mans	16	6	13	213
Lyon Grenoble	8	3	6	111
Lyon Clermont-Ferrand	13	5	10	165
Lyon Dijon	16	6	13	196
Lyon Chambery	8	2	6	108
Marseille Aix-en-Provence	3	1	1	33,2
Marseille Avignon	8	3	6	105
Marseille Toulon	5	1	4	66,3
Bordeaux Poitier	20	7	16	258
Bordeaux Pau	17	6	13	217
Toulouse Carcassonne	7	2	6	94,1
Montpellier Nime	4	1	3	56
Strasbourg Metz	13	4	10	165
Nantes Rennes	9	3	7	113
Nantes Anger	7	2	5	91,6
Secondary to secondary cities				
Lille Calais	9	3	7	111
Rennes Saint-Malo	5	2	4	69,5
Rennes Caen	15	5	12	186
Nice Cannes	3	1	2	33,1
Le-Havre Rouen	7	2	5	92,6
Nancy Metz	4	1	3	56,7
Tours Le-Mans	7	2	6	104
Tours Poitiers	8	3	6	112
Dijon Besancon	7	2	6	96

Prices are in euro. The return routes are not listed here. They belong to the same category as their pairs and share the same reference distance. However, a few routes' default price and price caps are different from their pairs.

Notes

We thank *Governance Analytics*, especially Bruno Chavas and Faten Amama, for their support of computer programming.