Bike-sharing services as an alternative in urban transportation: Evidence from a developing country*

Área 10 – Economia Regional e Urbana

Leonardo C. B. Cardoso†
Felipe de F. Silva‡

Abstract

We investigate how bike-sharing can contribute to urban mobility, combining an administrative database with more than twenty-five million bike-sharing trips with weather information, fuel prices, and data from Twitter. First, we ask whether there is a marginal substitution between private cars and bike-sharing driven by higher fuel prices. Second, we investigate if safety concerns deter a more extensive bike-sharing adoption. We find that a one standard deviation increase in gasoline prices is related to an increase of eight minutes of bike-sharing use per station per hour. Regarding the second question, an addition of one standard deviation in the number of negative tweets involving *bicycle* – our proxy for perceived safety – reduces bike-sharing use by more than one minute per station per hour.

Keywords: Bike-sharing, urban mobility, gasoline price, social media.

Resumo

Este artigo investiga como bicicletas compartilhadas (BC) podem contribuir para mobilidade urbana. Para tal propósito foram combinadoss dados de mais de vinte e cinco milhões de observações de viagens usando BC com informações sobre clima, preços de combustíveis e mídias sociais. Primeiro, é respondido se existe substituição marginal entre carros e BC motivada por preços de combustíveis mais altos. Depois, investiga-se como preocupações com segurança afetam o uso de BC. Os resultados apontam que um desvio-padrão a mais nos preços dos combustíveis aumenta o uso de BC em oito minutos por estaçao por hora. Aumentos de um desvio-padrão no risco percebido – medido pelo número de tweets negativos envolvendo o termo bicicleta – reduz o uso de BC em um minuto por estação por hora.

Palavras-chave: Bicicletas compartilhadas, mobilidade urbana, preço da gasolina, mídias sociais.

JEL Codes: R40, Q53, C50.

*The authors are grateful to Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) for funding this work. We also thank Renata Rabello for providing us permission to use the Tembici database.

†Departamento de Economia Rural, Programa de Pós-Graduação em Economia Aplicada, Universidade Federal de Viçosa, Viçosa, MG, Brazil. Corresponding author at leonardocardoso@ufv.br.

‡Agricultural Sciences Department, Agribusiness Program, Clemson University, Clemson, SC, U.S., fdsilva@clemson.edu.

1 Introduction

Douglas et al. (2011) draw a parallel between car ownership and tobacco consumption. Both goods can yield significant health harm, have a strong lobbying presence, and have consumption externalities. In Brazil, tobacco consumption reduced substantially between 2006 and 2017[1], but the transition to a less car-oriented urban transportation has been ignored for several reasons: (i) the aforementioned lobbying presence of the automotive industry in Brazil delays this transition; (ii) strategies to reduce car dependence are deeply unpopular in developing countries such as Brazil – car ownership has been seen as an out-of-poverty symbol; (iii) the low-quality alternatives, especially the public transportation, constrain individuals' ability to substitute car for other alternatives. Altogether, these factors along with other socioeconomic characteristics result in an elastic income elasticity[2] in this market, where minor income improvements are enough for rapid growth in car ownership. While the Gross Domestic Product (GDP) increased 26% in Brazil (World Bank, 2020) between 2006 and 2018, car ownership doubled (Denatran, 2022). These constitute an extra challenge for policymakers: *How to reduce car dependence when car ownership is quickly escalating and is expected to increase even more?*

In developed countries, several public policies attempted to reduce car-based mobility while incentivizing public transportation and other more sustainable urban transportation modes. Some of these policies put in place strategies that directly address car usage, such as carpool lanes (Cassidy et al., 2010) and highway and congestion tolls (Santos, 2005). Some indirectly increase the relative cost of cars by taxing fuel or reducing the commuting time gap between public transportation and individual cars (Liao et al., 2020). In this paper, we explore another strategy that accounts for an unusual transportation mode: bike-sharing.

In the last decade, bike-sharing stations skyrocketed in several countries, including Brazil. These stations are located in strategic places, e.g., next to metro and bus stations, where individuals can use rental bicycles for the first- and last-mile of their commute to work. This transport mode can provide sizable environmental benefit[3] and, indirectly, improve the health of its users. To better understand this alternative, we examine the determinants of bike-sharing use in Brazil and how to foster its use through public policies.

Nationally representative and detailed travel surveys including bicycles are usually not available in low- and middle-income countries. However, our paper takes advantage of an administrative database with more than twenty-five million georeferenced bike-sharing trips from Tembici, South America's leader in this micro-mobility with more than 70%

[1]The share of smoking adults in Brazil decreased from 15.7 to 10.1% due to tobacco control policies (Divino et al., 2021).

[2]De Negri (1998) find that income elasticity ranges from 1.1 to 1.5 in Brazil.

[3]This benefit can also be observed when these bikes are rented for touristic purposes, given that consumers would have taken a bus, Uber or a taxi, which emits greenhouse gases.

of this market. We focus on the bicycle rides between January 1, 2018, and December 31, 2019 of four major cities in Brazil: São Paulo, Rio de Janeiro, Salvador, and Porto Alegre. Around 24 million people live in this four cities. To estimate the determinants of bike-sharing usage and the potential effects of public policies, this data is merged with weather information, fuel prices, and social media data.

A large empirical literature focused on fuel demand suggests that substitution plays a major role in explaining how fuel price changes affect its consumption. These studies indicate that higher gasoline prices result in greater use of public transport (Lane, 2012), consumption of other fuels (diesel (Pock, 2010) and ethanol (Anderson, 2012)), and purchase of more fuel-efficient cars (Kayser, 2000). Cycling is not highlighted in this literature, especially because it represents a small share of daily trips in most countries[4]. In their defense, a substitution between gasoline and cycling is hard to be measured in fuel demand studies, where fuel consumption is the explained variable, due to the small share of daily trips using bikes. The *first research question* we explore in this paper is about a potential substitution between private cars and bike-sharing, but looking the other way around, with bike-sharing as an explained variable.

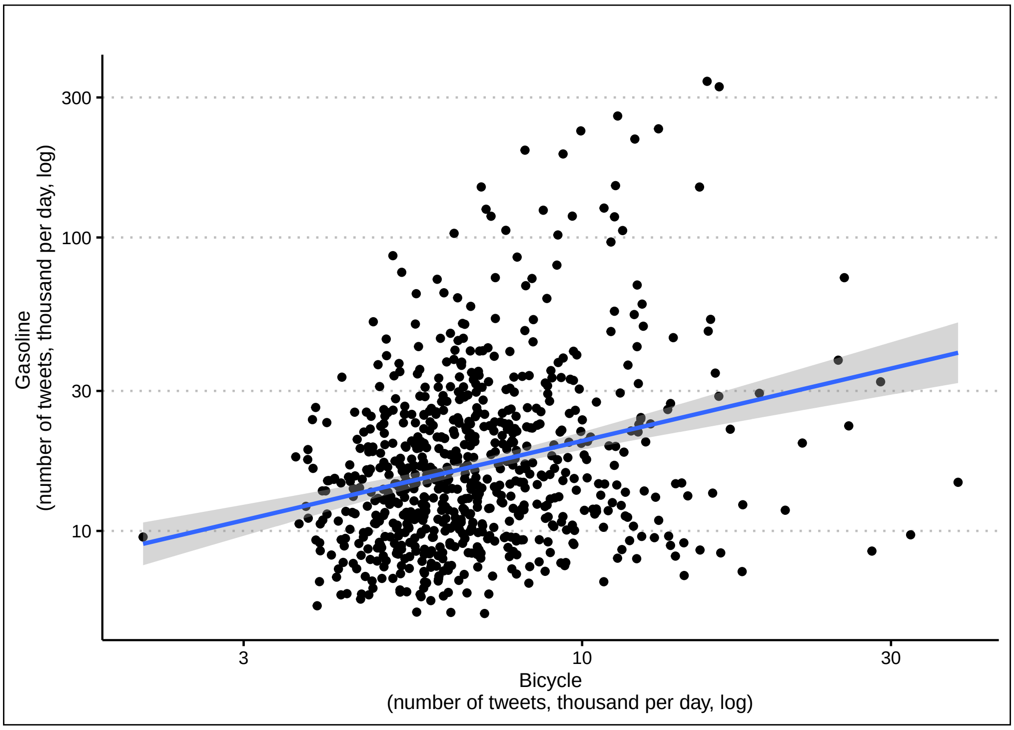
Pucher & Buehler (2008) indicate that countries with higher fuel prices have higher levels of cycling. This relationship might also manifest in Brazil. Our dataset is on bike-sharing instead of bicycle rides, which doesn't allow us to test Pucher & Buehler (2008) 's point. However, an unusual but correlated alternative is plotting the number of tweets with the term "gasoline" against those with the term "bicycle" (see Figure 1). This crude measure indicates a positive correlation between the two cited terms on social media in Brazil. Thus, there might also be a movement from private cars to bike-sharing during higher fuel prices. An interesting appeal to bike-sharing is that consumers don't need to buy a bicycle to ride; they only pay for the time they use, which implies no large capital investment. The substitution between car and bike is available immediately by renting a bicycle for 30 BRL (price in Month 2022) on monthly plans (roughly 6 USD using the foreign exchange rate of the Central Bank of Brazil on June-2022). So, specifically, we estimate the effect of marginal increases in the cost of private car trips, measured in terms of gasoline prices, on bike-sharing use. Results from a two-part model indicated that higher gasoline prices increase bike-sharing use after controlling for several observables, such as weather, dock station, month, day, and hour fixed effects.

To account for the potential issue that gasoline price is also measuring omitted time-varying variables, we delve deeper into the relationship between fuel prices and bike-sharing exploiting an exogenous rise in gasoline prices caused by truck drivers' strikers on May 21, 2018. The gasoline prices increased seven percent in Brazil because of the

[4]The share of bicycle trips in Brazil is close to the US, which is around one percent. In contrast, the Netherlands and Germany have higher shares; more than 25% of their citizens' daily trips are by bicycles (see Goel et al. (2022) for a comprehensive ranking).

strike. Using a regression discontinuity approach (RDD), we estimate how bike-sharing consumers respond to increases in gasoline prices. Results remain unchanged: increases in gasoline prices increase bike-sharing use. The results indicate that a rise in 10% gasoline price increases is related to increases of 30% in the number of trips and 40% in the minutes of use. Barriola (2021) and He et al. (2020) provide an initial analysis of the relationship between fuel prices and bike-sharing use, and find that higher fuel prices result in more bike-sharing use. Differently, our paper uses a quasi-experimental method to estimate the effect of fuel prices on bike-sharing use in the Brazilian market.

Figure 1: Gasoline versus Bicycle terms on Twitter.



Notes: The figure shows a scatter plot of numbers of tweets with the "gasoline" term (vertical axis) against tweets with the "bicycle" term (horizontal axis), both in logarithm scale. We collected tweets in Portuguese between January 1, 2018, and December 31, 2019, using the library developed by Barrie & Ho (2021). The best linear fit line is in blue. The correlation between the two variables is 0.22.

Higher fuel prices can increase bike-sharing use but safety concerns might discourage cycling and bike-sharing. In the Europe Union, more than nineteen thousand cyclists were killed between 2010 and 2018, representing eight percent of road deaths (Adminaité-Fodor & Jost, 2020). In Brazil, cyclists have relatively more traffic-related deaths. While trips using bicycles are only one percent, their deaths represent five percent of the total deaths on roads in Brazil (CET, 2021). It explains why cyclists, jointly with pedestrians, are classified as *vulnerable road users*. Nevertheless, correctly measuring how safety concerns impact cycling is not an easy task. The usual proxies for individuals' safety concerns are measuring how cyclists interact with motorized vehicles, e.g., the presence of calming traffic streets, pedestrians and separated bike lanes, road speed limit, lane width, and the volume of motor vehicles (Klobucar & Fricker, 2007). However, all these variables shaped long-run perceptions of cyclists.

The *second contribution* of this paper is to introduce a short-run perceived risk measure, and investigate how it relates to bike-sharing use. To measure individuals' safety concerns, we use Twitter data. We observe that new trends on Twitter emerge when tragedy involving bicycles occurs. Large spikes in the number of tweets combining *bicycle* and negative words (*accident* and *death*) occur when there was massive media coverage of road traffic accidents involving cyclists. The key to our approach is that media coverage does not change risk but rather changes how users perceive risk in the short-run. So, based on a broad literature on how media exposure can affect social outcomes (Kearney & Levine, 2015; Gentzkow & Shapiro, 2008; Jensen & Oster, 2009) we argue that tweets combining the terms *bicycle* and *accident* and *bicycle* and *death* can be used as proxy of how safe bike-sharing users are feeling in the short-run. After controlling for weather, dock station, month, day, and hour fixed effects, we find that bike-sharing usage is negatively affected by higher levels of perceived risk.

The rest of this paper is organized as follows. Section 2 describes our data sources and empirical strategy. Section 3 contains the main results and discussion on their implication for public policy. Finally, the last section concludes the paper, discusses the future agenda as well as presents the limitations of our work.

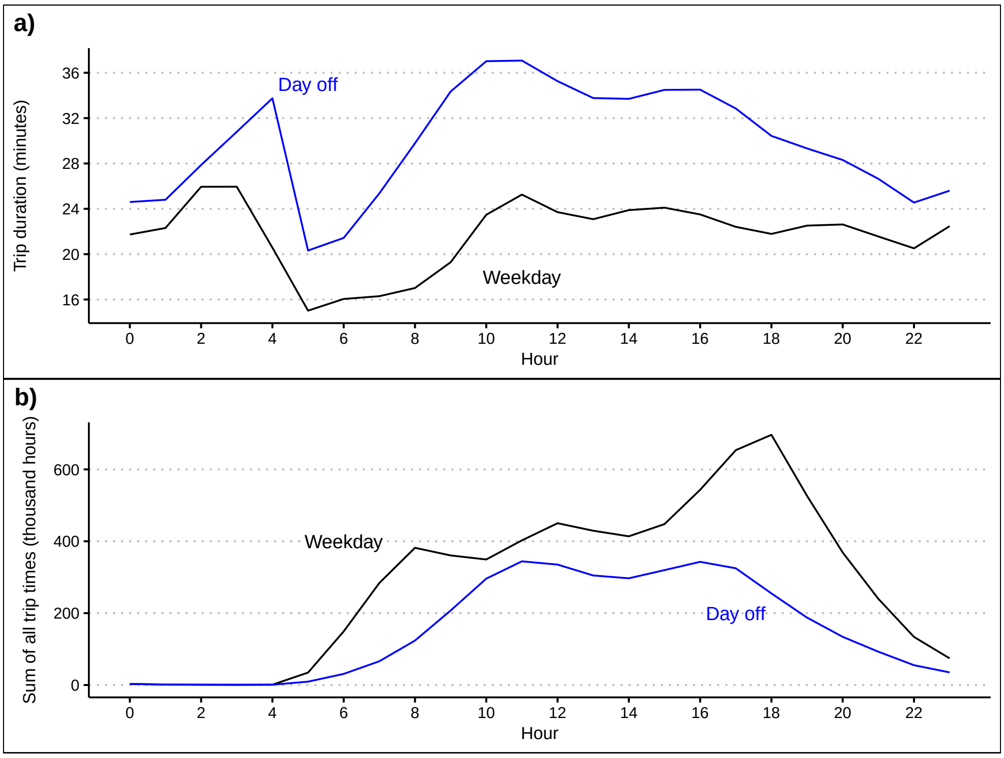
2 Method

2.1 Data

The data on the bike-sharing trips is from Tembici and includes information such as the exact time and station when trips started and ended, which allows us to geocode trips as well as compute their duration. The database contains two of the biggest cities in Brazil (São Paulo-SP and Rio de Janeiro-RJ) and two major state capitals (Salvador-BA and Porto Alegre-RS). This initial database allows an understanding of some consumption patterns. For example, trips tend to be longer during days off (Saturday, Sunday, and Holiday) compared to weekdays (see Figure 2a). Looking at the intra-day time differences, the beginning of the morning has the fastest trips (Figure 2a). The sum of all trip durations (Figure 2b) shows an overall more intensive use of bike-sharing during weekdays, with the peak exactly during rush hours in Brazil (between 17:00 and 19:00).

We combined the bike-sharing trips with a temperature, precipitation, and humidity database from the Instituto Nacional de Meteorologia do Brasil (INMET), where variables vary by city, date, and hour. Cities with higher precipitation levels are not cycling-intensive (Hong et al., 2020). This is confirmed in our database; simple descriptive statistics indicate a reduction of 51% in bike-sharing trips on days with any precipitation

Figure 2: Trip duration and bike-sharing use by hours.



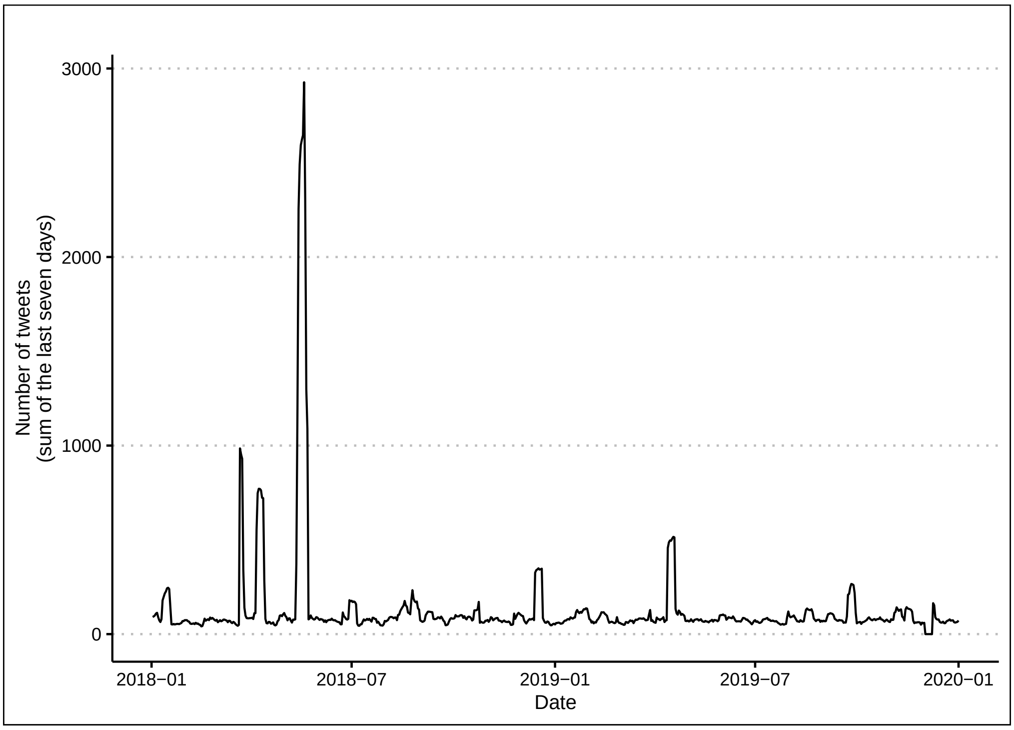
Notes: "Day off" and "Weekday" split sample between days off (Saturday, Sunday, and Holiday) and workdays (every other day different from days off). In Figure 2a we have the average trip duration in each hour for weekdays and days off. In Figure 2b we have the sum of all trip durations in each hour also for weekdays and days off.

compared with days with no precipitation. There are no clear trends regarding other weather conditions (temperature and humidity).

We merged trip and weather databases with data from Twitter. Each message on Twitter, a tweet, is limited to 140 characters. It is relatively easy to access recent tweets – messages from the last seven days – using packages in RStudio and Stata. However, as mentioned by Kearney & Levine (2015), accessing tweets posted more than a week ago was not a free asset. It was done previously through a third-party vendor. Fortunately, very recently, a new library on RStudio – academictwitteR, developed by Barrie & Ho (2021) – allows us to have access to older tweets for free.

There is a limitation in the geographic detail of tweets – only a few users identify where they live. As we used the number of tweets with bicycle and accident-related terms as a measure of perceived risk for the whole country (see Figure 3), a drawback is the possibility of collecting tweets from other Portuguese-speaking countries (Portugal and Mozambique, for example). However, Brazil is the fourth country in the number of Twitter users (twitters), with nineteen million users. In the ranking of twenty countries with the most Twitter users, there is no other country that speaks Portuguese. The last country among these top 20 is Colombia, with four million Twitter users in Jan-2022 (Dixon, 2022). Thus, by collecting tweets in Portuguese, there is a high probability that we are actually collecting Brazilian tweets. In the analysis, we use the sum of the last seven days because we believe the news has an impact on the decision to rent bike-sharing during the whole week.

Figure 3: The number of tweets containing *bicycle* and *accident* or *bicycle* and *death* in the last seven days.



Note: Tweets were collected between Jan 01, 2018, and Dec 31, 2019, using terms in Portuguese, and included re-tweets.

The last added layer of data is the gasoline prices from the Sistema de Levantamento de Preços of the Agência Nacional do Petróleo, Gás Natural e Biocombustíveis (ANP, 2022). ANP provides weekly prices for all four cities in our database. In Brazil, there is no pure gasoline sold at pumps; car owners usually buy E100 (100% of hydrated ethanol), E27 (a blend with 73% of gasoline and 27% of anhydrous ethanol - the Gasoline-C), or E25 (premium gasoline with a slightly lower share of anhydrous in the blend, 25%). The price used is the pump price of Gasoline-C.

For our empirical strategy, we aggregated the trip level to a station-hour panel data to generate the variables: number of trips per station per hour (henceforth, number of trips) and number of minutes per station per hour (henceforth, minutes of use). To be consistent, all other variables were collapsed accordingly. The final sample used in the main regressions has more than eight million observations. The summary of statistics is in Table 1. On average, a trip lasts 33 minutes and there were 1.33 trips per hour and station.

Table 1 also displays the control variables used in the regression analysis. Station (*id*) represents each dock station where bicycles can be rented from the Tembici company (we

Table 1: Descriptive Statistics

	Observations	Mean	Std	Min	Max
Trip-panel					
Duration (minutes)	26,094,893	24.63	25.93	0	600
Precipitation (mm)	24,266,793	0.09	0.96	0	71.6
Temperature (Celcius)	24,265,868	23.81	3.91	2.30	38.0
Gender	6,792,435	1.73	0.44	1	3
Station hour panel					
Number of trips	8,569,800	1.33	4.87	0	472
Minutes of use	8,569,800	33.64	123.99	0	28604.43
Station (id)	8,569,800	518	298.78	1	1035
Gasoline Price (BRL)	8,540,912	4.49	0.36	3.74	5.26
Hour	8,478,671	13.50	4.60	6	21
Temperature (0-4)	8,556,117	2.36	1.15	0	4
Day off	8,556,117	0.32	0.46	0	1
Month	8,478,671	5.80	3.21	1	12
City (1-4)	8,569,800	2.51	0.76	1	4
Day off	8,556,117	0.32	0.46	0	1
Number of tweets	8,479,620	1.96	6.86	0	100
Year	8,556,117	2018.35	0.47	2018	2019

Notes: There were a few trips with more than ten hours, we dropped those. In the station hour panel, we consider only trips between 6:00 am and 21:00. The number of tweets results from the sum of tweets in the last seven days – we used a min-max normalization (0 for the minimum and 100 for the maximum).

used it as fixed effects: 1 for each unique *id* and zero otherwise). We convert the variable *Hour* in a fixed effect, a dummy variable for each hour of the day, regardless of the minutes; we used 15 fixed effects rather than 23 because we considered trips between 6:00 and 21:00 in the station-hour panel. *Month* is also used as a fixed effect in the regression analysis, we use eleven fixed effects, and *year* is a dummy to differentiate between 2018 and 2019. *Temperature* is a continuous variable in the trip-panel data, but a categorical variable in the hour-station panel. Then, we constructed four categorical variables (0 to 4), where it is equal to zero if the temperature is lower than 18 degrees Celsius, one if it is between 18 and 22, two for values between 22 and 25, three for temperatures between 25 and 28, four for those higher than 28 degrees Celsius. This variable was used to build dummy variables used in the regression.

Unfortunately, we only have gender information for a quarter of the sample. The share of females in users is higher than 28% in all cities, except São Paulo, which has only 18% of females among users. The rest of the users are males[5].

[5]Other genders have a residual value lower than 0.1% of users in all cities.

2.2 Empirical Strategy

2.2.1 Two-part model (TPM)

The structure of the two dependent variables (number of trips or minutes of use) makes ordinary least squares (OLS) questionable because of the mass of zeros that naturally appear in the construction of a station-hour panel – a lot of stations have no trips in some hours. The use of a Heckman selection model is questionable as well because our zeros are not censored values as in a Heckman model. In our panel, zeros are true values. To overcome the zero-issue, we used a two-part model[6].

Two-part models have been used by meteorologists to predict rainfall (Pegram et al., 2009) and by economics to predict health care expenditures (Deb & Trivedi, 2002). The basic idea is to have the first stage considering the participation (positive-versus-zero outcome) modeled by a parametric binary probability model. And, conditional on participation, a second stage fit for the positive outcome (Mullahy, 1998). We used a logit distribution for the first stage and OLS for the second one. The first stage can be represented by Equation 1 and the second stage by Equation 2 (we are following Belotti's notation[7]).

$$\phi(y > 0) = Pr(y > 0 \mid x) = F(x\delta) \tag{1}$$

$$\phi(y \mid y > 0, x) = G(x\gamma) \tag{2}$$

where x is a vector of explanatory variables – we used the same vector of explanatory variables (x) for Equations 1 and 2. δ is the vector of parameters to be estimated in the first stage (Equation 1) and γ is the vector of parameters to be estimated in the second stage (Equation 2). F and G are density functions. The vector of explanatory variables has the nominal gasoline prices in BRL (level) and the sum of tweets from the last seven days as variables of interest, and several controls. The gasoline price is a proxy for private car trip cost and the tweets are used as a proxy for perceived risk. All other variables in Table 1 are controls (station, hour, temperature, precipitation day off, month, city and year).

2.2.2 Regression Discontinuity Design (RDD)

To estimate the causal effect of fuel prices on the variables of interest (y = number of trips or minutes of use), we explore an exogenous increase in gasoline prices caused by the truckers' strike on May 21, 2018. It allowed us to measure the effects of gasoline prices as a local randomized experiment. The strike lasted nine days and gasoline prices increased by seven percent compared to the days before and after the strike. The national average

[6] TPM is a "hurdle" model but for continuous data (Belotti et al., 2015).

[7] See Belotti et al., 2015 for more details.

of gasoline prices increased from 4.28 BRL (May 19, 2018) to 4.61 BRL (June 02, 2018) and remained around 4.60 BRL until the middle of June. For the analysis, we constructed two other panel datasets, one accounting for 14 days, and another accounting for 21 days before and after the strike following Equation 3.

$$Y_{it}^{n} = \beta * 1(discontinuity) + \alpha_i + \gamma_t + \theta.weather + \epsilon_{it} \tag{3}$$

where discontinuity = 0 if date < May 21, 2018; and discontinuity = 1 after May 21, 2018. During the strike, the effects of higher gasoline prices could be mixed with a concern of fuel shortage – the transport between refineries and gas stations is mostly done by trucks in Brazil. In this way, consumers may worry about running out of gasoline for an emergency, and decreasing gasoline consumption not exclusively due to higher prices but also to shortage concerns. Thus, we tested another specification comparing pre- and post-strike but excluding the nine days during the strike.

3 Analysis and Results

3.1 Two-part model estimates

We start with the analysis of the two-part model. The results in Table 2 indicate that the costs of private car trips, measured by gasoline prices, increase bike-sharing use. The elasticity of bike-sharing use with respect to gasoline prices is between three (when analyzing the minutes of use) and four (when analyzing the number of trips). It implies that increases of 10% in gasoline prices would lead to increases of 30% in the number of trips and 40% in the minutes of use. The methodology used here also allow us to estimate the increase in the number of rides in each station per hour due to a change in the gasoline prices. We are most interested in the case when there wasn't any rental, in other words, answering the question *How many trips would a quiet station observe in a given hour for an increase of one standard deviation on gasoline price?* We find that gasoline prices increase the probability[8] of renting the first bicycle in a station in a given hour that previously had no use as well as in station in a given hour that already have been used. One extra standard deviation in gasoline prices (0.36 BRL) would lead to 0.44 more trips per station per hour. In terms of minutes, one additional standard deviation in gasoline price is related to an increase of almost eight minutes on the average minutes of use per station per hour.

The estimated bike-sharing elasticity with respect to gasoline price might seem large

[8]Note that the first stage, Logit estimation, is the same for both models (different dependent variables). It is because when a station does not have any trip, it also does not have any minute of use, and when a station has zero minutes of use, it has no trip. Also, we include the same independent variables in both models.

for readers suited to fuel demand literature. Even in Brazil, where gasoline price elasticity is larger than in developed countries, estimates are rarely larger than one[9]. However, bike-sharing consumption is more volatile than gasoline consumption. Bike-sharing use increased 50% between the first and the second semester of 2019 (Tembici, 2022). Such change between two semesters is not observed in the gasoline demand. For instance, fuel sales reduced less than 18% between the first semester before the coronavirus pandemic in 2019 and the first during the outbreak in 2020 (ANP, 2022). Also, keep in mind, that we are estimating the cross-price elasticity – how the demand for bike-sharing services changes with changes in the prices of a substitute (cost of using a private car). These elasticities tend to be larger.

Regarding the impact of our perceived risk variable, the number of tweets combining bicycle and negative words shows a negative impact on minutes of use in both stages (column 2, Table 2). Even though we find that perceived risk explains whether a station in a given hour has a trip, we do not find a significant effect on the number of trips (column 1, Table 2). The perceived risk seems to have a clearer impact on minutes used (column 2, Table 2) than on the number of trips. In the second stage, for example, the elasticity of minutes of use with respect to the number of tweets is -1.29, indicating that more 10% in the number of negative tweets is related to a reduction of 12.9% in minutes of use. One extra standard deviation in the number of tweets (6.89) is related to 1.1 more minutes of use per station per hour.

Some other controls have a significant role in the probability of renting a bicycle. The number of trips and minutes of use increased during day offs (Saturday, Sunday, and Holidays) and on days with no precipitation. Regarding temperature, we have a higher probability of renting a bicycle when the temperature is between 18 and 25 Celsius compared to temperatures higher than 25 Celsius.

[9]See Santos (2013) and Cardoso et al. (2019).

Table 2: Two-part model estimates

	(1) Number of trips	(2) Minutes of use
Logit		
Gasoline price	1.23***	1.23***
	(272.87)	(272.87)
	$\varepsilon = 4.13$	$\varepsilon = 4.13$
Number of tweets	-0.0189***	-0.0189***
	(-11.31)	(-11.31)
	$\varepsilon = -0.03$	$\varepsilon = -0.03$
OLS		
Gasoline price	1.24***	22.22***
	(35.17)	(24.97)
	$\varepsilon = 4.16$	$\varepsilon = 2.96$
Number of tweets	0.0013	-0.160***
	(1.38)	(-6.55)
	$\varepsilon = 0.002$	$\varepsilon = -1.29$
Observations	8,458,331	8,458,331

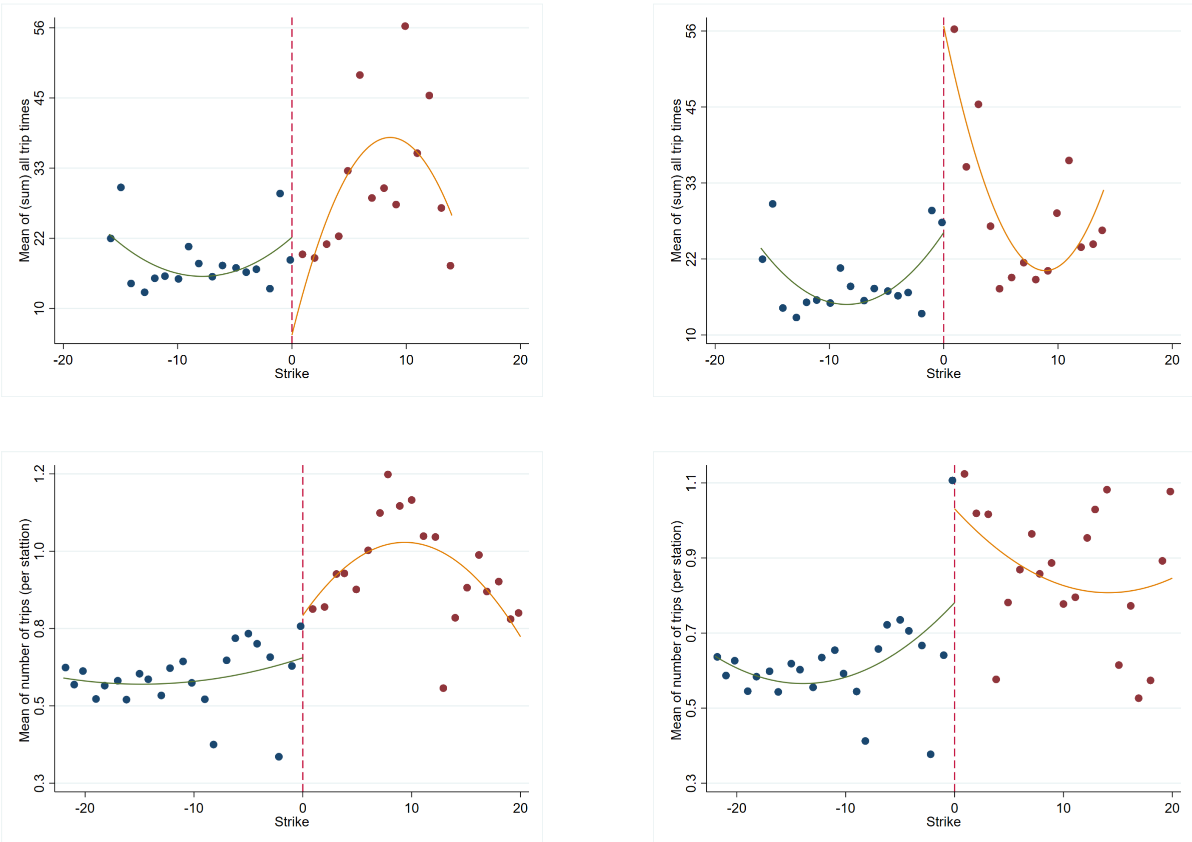
Notes: t statistics in parentheses. $*p < 0.05$, $**p < 0.01$, $***p < 0.001$. The regressions reported are based on a station-hour panel. All regressions have the following fixed effects: station, hour, day, month, year, and temperature (five categories). Elasticities (ε) were calculated at means using the Delta method.

3.2 Regression Discontinuity Design (RDD) estimates

In this section, we will focus on data panels constructed to estimate the impact of fuel prices on variables of interest given an exogenous change - a truck strike in May of 2018. The gasoline prices post-strike period was seven percent higher than before the strike. Figure 4 shows bike-sharing use for the pre- and post-strike periods. It provides some visual evidence of our discontinuity design. It shows an increase in the number of trips and minutes of use in the post-strike period in all four panels. Panels A reveals how average minutes of use behave 14 days pre and post-strike, including days during the strike. Panel B displays the same information but excludes strike days. Panels C and D show how the mean number of trips behaves in the 21 days before and after the strike as well, including and excluding the nine days of strike, respectively.

In Table 3 we presented the regression discontinuity estimates for the effect of gasoline prices in these four different scenarios and for the other four. All the results point to elasticities around three and four – in line with our two-part model estimates. The exceptions are the minutes of use in a fifteen days window (columns 3 and 4 in Table 3), which show elasticities higher than six.

Figure 4: The number of trips and sum of minutes of bike-sharing use before and after the truck drivers' strike.



Notes: Dots are the average sum of hours of bike-sharing use per station per hour. The fitted curves are a separated quadratic polynomial for the pre-strike period (green curve), and another quadratic polynomial for the post-period (orange curve). The horizontal axis represents the number of days before and after the strike. Panels A and C include days during the strike as post-period. Panels B and D exclude the nine days of strike.

Table 3: Regression Discontinuity Design (RDD) estimates for different sub-samples exploring an exogenous change - a truck strike on May of 2018.

Panel A = Strike days not included				
	(1)	(2)	(3)	(4)
	Number of trips	Number of trips	Minutes of use	Minutes of use
Discontinuity	0.40***	0.30***	18.45***	11.25***
	(37.73)	(40.54)	(43.13)	(39.74)
	$\varepsilon = 4.17$	$\varepsilon = 3.12$	$\varepsilon = 7.62$	$\varepsilon = 4.65$
Window	14 days	21 days	14 days	21 days
Strike days included	No	No	No	No
Observations	340024	534149	340024	534149
Panel B = Strike days included				
	(5)	(6)	(7)	(8)
	Number of trips	Number of trips	Minutes of use	Minutes of use
Discontinuity	0.37***	0.31***	14.76***	11.02***
	(48.11)	(49.14)	(51.86)	(46.45)
	$\varepsilon = 3.85$	$\varepsilon = 3.23$	$\varepsilon = 6.09$	$\varepsilon = 4.55$
Window	14 days	21 days	14 days	21 days
Strike days included	Yes	Yes	Yes	Yes
Observations	486099	680224	486099	680224

Notes: t statistics in parentheses. $*p < 0.05$, $**p < 0.01$, $***p < 0.001$. The regressions reported are based on a station-hour panel. All regressions have the following fixed effects: station, hour, day, month, year, and temperature (five categories). Elasticities (ε) at point: $E = \beta/Y(y \mid discontinuity = 0) * \Delta P)$. Panel A drops the nine-day strike period from the sample. Panel B includes these nine days. Window stands for the number of days pre- and post-strike.

4 Conclusion and Policy Implications

In this paper, we investigate how bike-sharing can contribute to urban mobility, specifically identifying which factors can foster bike-sharing use while reducing car dependence. While several studies investigate cycling and bike-sharing services in developed countries, this is the first study focusing on a large developing country while discussing the relevance of fuel consumption. *Why focus on fuel consumption?* There is extensive literature on the impacts of fuel consumption on the environment and population health. Bike-sharing services have a unique appeal because while it decreases the environmental effect (GHG emission), they also provide a healthy activity. One may say that Brazil is not a traditional cycling society (it lacks the culture of cycling), which is a crucial factor in higher cycling levels. However, if cities can learn with the right incentives - and they do, Seville in Spain went from one to six percent of bicycle trips within four years (Geller & Marqués, 2021) - countries can learn as well. In the case of Brazil, demand for these services increased by 50% in one year.

The first contribution of this paper is that we find that there is a marginal substitution between private car trips and bike-sharing led by higher fuel prices in Brazil. One standard deviation change in gasoline prices would increase bike-sharing use by eight minutes per station per hour. A whole year with gasoline prices one standard deviation higher than usual would lead to more than five million kilometers using bicycles. This estimate is considering only Tembici stations in the four cities of our database (São Paulo, Rio de Janeiro, Porto Alegre, and Salvador)[10]. If users had each a private car to travel these additional five million kilometers (ten kilometers per liter of gasoline), additional emissions would be more than one million tons of carbon equivalent. These results were confirmed when exploring an exogenous change, truck strick in May of 2018, using a Regression Discontinuity Design approach. In conclusion, the positive relationship between fuel prices and bike-sharing services is robust to a variety of statistical specifications and samples, which reassures any policy analysis made here.

A second significant contribution was related to how safety concerns affect demand for these services. We used the number of tweets with bicycle combined with negative terms (*accident* and *death*) as a proxy for a short-run perceived risk measure. Our proxy is negatively related to the number of trips and minutes of use in bike-sharing. So, there is a reduction in bike-sharing use when short-run perceived risk increases. In the paper, we report results only on the number of tweets in the last seven days as a proxy for safety concerns. We have also considered a variable capturing *search intensity* by Google

[10]It is a lower bound estimate using the following information: a) there are roughly 600 stations in the four cities in our sample; b) we consider rides between 6:00 and 21:00 (15 hours). X = 600*365*15*8 = 26280000 minutes = 438000 hours. Then we translate it to kilometers using the geographic coordinates of users who return bicycles at a different station than the ones they picked up. On average, they travel 12 kilometers in one hour.

Trends, and the results are similar to those reported here. However, after analyzing a few different specifications, we notice that how the variable *perceived risk* is built affects the results; specifically, the lag and summation length (in days) of the tweets play a role in the impact of this perceived risk variable on demand for this services[11]. It is an indication that more work is needed to estimate this effect correctly. As this is an innovative way to measure perceived risk, we did not find literature that supports it.

The results obtained from tweets are an indication of how short-run events can affect bike-sharing services in these cities. An analysis of the effect of long-run changes is missing here and in the literature for Brazil. However, as short-run events (tweets) are a reflection of current infrastructure, among other factors, our results suggest that as the transition to a less car-dependent urban transportation or more bike-friendly occurs, we could expect an increase in the use of these services given that safety (or at least how it is perceived) is expected to increase (implies a decrease on the above mentioned tweets).

In this version, we studied the impact of two simple short-run factors that can foster bike-sharing adoption, fuel prices and perceived safety. In the subsequent versions, we will also investigate how these factors (fuel prices and safety concerns) can affect long-run decisions related to urban mobility, in this case, bike-sharing services. As well as add other factors that would shape long-run decisions, e.g., bicycle infrastructure and connectivity levels between different transportation modes. As we have access to high-frequency data for bike-sharing, combining it with traffic databases would be interesting to identify whether traffic jams affect the demand for bike-sharing services. This could answer whether users are more likely to rent bicycles at times of greater traffic jams.

[11]Note that the results on fuel prices remain unchanged.

References

Adminaité-Fodor, D. & Jost, G. (2020), 'How safe is walking and cycling in Europe? pin flash report 38.'.

Anderson, S. T. (2012), 'The demand for ethanol as a gasoline substitute', *Journal of Environmental Economics and Management* **63**(2), 151–168.

ANP (2022), *Sistema de Levantamento de Preços (SLP) da Agência Nacional do Petróleo, Gás Natural e Biocombustíveis (ANP)*. URL http://www.anp.gov.br/precos-e-defesa-da-concorrencia/precos/levantamento-de-precos.

Barrie, C. & Ho, J. C.-t. (2021), 'academictwitter: an r package to access the twitter academic research product track v2 api endpoint', *Journal of Open Source Software* **6**(62), 3272.

Barriola, X. (2021), 'The impact of gasoline shortages on public bike-sharing systems', *Academia Revista Latinoamericana de Administración*.

Belotti, F., Deb, P., Manning, W. G. & Norton, E. C. (2015), 'twopm: Two-part models', *The Stata Journal* **15**(1), 3–20.

Cardoso, L. C., Bittencourt, M. V., Litt, W. H. & Irwin, E. G. (2019), 'Biofuels policies and fuel demand elasticities in Brazil', *Energy policy* **128**, 296–305.

Cassidy, M. J., Jang, K. & Daganzo, C. F. (2010), 'The smoothing effect of carpool lanes on freeway bottlenecks', *Transportation research part A: policy and practice* **44**(2), 65–75.

CET (2021), 'Prefeitura da Cidade de São Paulo', *Companhia de Engenharia de Tráfego, Secretaria Municipal de Mobilidade e Trânsito*. URL http://www.cetsp.com.br/consultas/bicicleta.

De Negri, J. A. (1998), 'Elasticidade-renda e elasticidade-preço da demanda de automóveis no brasil'.

Deb, P. & Trivedi, P. K. (2002), 'The structure of demand for health care: latent class versus two-part models', *Journal of health economics* **21**(4), 601–625.

Denatran (2022), 'Frota de Veículos - 2022', *Departamento Nacional de Trânsito, Ministério da Infraestrutura, Governo Federal*. URL https://www.gov.br/infraestrutura/pt-br/assuntos/transito/conteudo-Senatran/frota-de-veiculos-2022.

Divino, J. A., Ehrl, P., Candido, O. & Valadao, M. A. P. (2021), 'Extended cost–benefit analysis of tobacco taxation in Brazil', *Tobacco Control*.

Dixon (2022), 'Leading countries based on number of Twitter users as of January 2022', *Statista*. URL https://www.statista.com/statistics/242606/number-of-active-twitter-users-in-selected-countries/.

Douglas, M. J., Watkins, S. J., Gorman, D. R. & Higgins, M. (2011), 'Are cars the new tobacco?', *Journal of Public Health* **33**(2), 160–169.

Geller, R. & Marqués, R. (2021), '19 implementation of pro-bike policies in portland and seville', *Cycling for Sustainable Cities* p. 74.

Gentzkow, M. & Shapiro, J. M. (2008), 'Competition and truth in the market for news', *Journal of Economic perspectives* **22**(2), 133–154.

Goel, R., Oyebode, O., Foley, L., Tatah, L., Millett, C. & Woodcock, J. (2022), 'Gender differences in active travel in major cities across the world', *Transportation* pp. 1–17.

He, P., Zou, Z., Zhang, Y. & Baiocchi, G. (2020), 'Boosting the eco-friendly sharing economy: the effect of gasoline prices on bikeshare ridership in three us metropolises', *Environmental Research Letters* **15**(11), 114021.

Hong, J., McArthur, D. P. & Stewart, J. L. (2020), 'Can providing safe cycling infrastructure encourage people to cycle more when it rains? the use of crowdsourced cycling data (strava)', *Transportation research part A: policy and practice* **133**, 109–121.

Jensen, R. & Oster, E. (2009), 'The power of tv: Cable television and women's status in india', *The Quarterly Journal of Economics* **124**(3), 1057–1094.

Kayser, H. A. (2000), 'Gasoline demand and car choice: estimating gasoline demand using household information', *Energy economics* **22**(3), 331–348.

Kearney, M. S. & Levine, P. B. (2015), 'Media influences on social outcomes: The impact of mtv's 16 and pregnant on teen childbearing', *American Economic Review* **105**(12), 3597–3632.

Klobucar, M. S. & Fricker, J. D. (2007), 'Network evaluation tool to improve real and perceived bicycle safety', *Transportation Research Record* **2031**(1), 25–33.

Lane, B. W. (2012), 'A time-series analysis of gasoline prices and public transportation in us metropolitan areas', *Journal of Transport Geography* **22**, 221–235.

Liao, Y., Gil, J., Pereira, R. H., Yeh, S. & Verendel, V. (2020), 'Disparities in travel times between car and transit: Spatiotemporal patterns in cities', *Scientific reports* **10**(1), 1–12.

Mullahy, J. (1998), 'Much ado about two: reconsidering retransformation and the two-part model in health econometrics', *Journal of health economics* **17**(3), 247–281.

Pegram, G. G. et al. (2009), 'A nested multisite daily rainfall stochastic generation model', *Journal of Hydrology* **371**(1-4), 142–153.

Pock, M. (2010), 'Gasoline demand in Europe: New insights', *Energy Economics* **32**(1), 54–62.

Pucher, J. & Buehler, R. (2008), 'Making cycling irresistible: Lessons from the netherlands, denmark and germany', *Transport Reviews* **28**(4), 495–528. URL https://doi.org/10.1080/01441640701806612.

Santos, G. (2005), 'Urban congestion charging: a comparison between london and singapore', *Transport Reviews* **25**(5), 511–534.

Santos, G. F. (2013), 'Fuel demand in Brazil in a dynamic panel data approach', *Energy Economics* **36**, 229–240.

Tembici (2022), 'A atuação da Tembici na América Latina', *Micromobilidade ao sul global*. URL https://www.mobilize.org.br/estudos/483/micromobilidade-no-sul-global.html.

World Bank (2020), 'GDP per capita (current US$) - Brazil', *World Development Indicators, The World Bank Group*. URL https://data.worldbank.org/indicator/NY.GDP.PCAP.CD?locations=BR.