



**HAL**  
open science

# A comparative study of gravity models to assess the evolution of urban bicycle mobility

Bertrand Jouve, Paul Rochet, Mohamadou Salifou

## ► To cite this version:

Bertrand Jouve, Paul Rochet, Mohamadou Salifou. A comparative study of gravity models to assess the evolution of urban bicycle mobility. 2024. hal-04714012

**HAL Id: hal-04714012**

**<https://hal.science/hal-04714012v1>**

Preprint submitted on 30 Sep 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# A comparative study of gravity models to assess the evolution of urban bicycle mobility

Bertrand Jouve<sup>1</sup>, Paul Rochet<sup>2</sup>, and Mohamadou Salifou<sup>1,2</sup>

<sup>1</sup>LISST, University of Toulouse, CNRS, University of Toulouse Jean Jaurès

<sup>2</sup>Federation ENAC ISAE-SUPAERO ONERA, University of Toulouse, France.

## Abstract

Gravity models are extensively used tools to understand mobility, providing an approximation of the volume of a flow as a function of distance and local population density. Specific gravity models have been designed for studying bike sharing systems (BSS) in urban areas. In this paper, we show how robustness to scale change in data pre-processing can be used as a criterion to select the best model for BSS usage, in addition to the usual criteria of quality of fit. In a second stage, the study the recent evolution of BSS use in large cities via a dynamical application of gravity models over rolling time windows. Exploiting the good interpretability of gravity models, this methodology produces a time series for each parameter, that can be analysed to study the evolution of local policies and users habits.

## 1 Introduction

Due to its ease of use, reduced environmental impact and health benefits, bicycle has become an increasingly popular mode of transportation in urban areas [37]. To promote bicycle use [8, 25, 15, 24], public authorities have implemented Bike-Sharing systems (BSS), with obvious desirable side effects such as the reduction of traffic and air pollution. The COVID-19 pandemic has clearly boosted the popularity of cycling in the city [4]. The undeniable success of BSS all over the world has naturally led to an ever increasing interest from the scientific community. A general overview on the topic can be found for instance in [10, 33, 11, 22].

Most models of urban mobility concentrate on linking two stations, say  $i$  and  $j$  with respect to distance and additional geographical and/or social factors. Gravity models, inspired by Newtonian mechanics, provide an elegant way to study these interactions. These models have been widely used in the literature to study mobility in general, within or among cities [3, 14, 18], and can be made more complex to better fit the observed flows of individuals by adding external variables related, for example, to land use or transport [27].

Since a model hardly ever produces a perfect fit on the data, the implicit assumption is the existence of a stochastic residual term which measures the deviation from each data point to the model prediction, and only depends on external unobserved factors (hence the resort to randomness). Neglecting to study the residual terms in the model can be problematic however, as their distributions should typically dictate the inference method. Based on criteria of simplicity, accuracy and

robustness to scale change, Poisson distributed gravity models [12] with an exponential relationship to distance emerge as the best contender for studying BSS data. We implement these models on datasets from different cities’ bike-sharing systems (BSS), with an emphasis on the French city of Toulouse and the American cities of Chicago and Washington. By analysing these data in a dynamic fashion over the four-year period between January 2019 and December 2022, we are able to study the temporal evolution of the BSS practice over different time windows. In particular, we highlight the weekly adaptation of the population habits to sudden environmental or societal constraints, an infamous example being the Covid-19 crisis, the impact of which is already well documented both in a broad context [29, 34, 2, 1, 6] and in specific locations such as New-York [31, 36, 23], Chicago [16], Lisbon [32], London [20], Seoul [30], Singapore [28], Zurich [19] or Beijing [5, 26] to name a few.

The geographical scale of the data is known to strongly impact the outcome of gravity models, to the extent that modified versions are typically more suited to inter-cities data [35, 38], compared to urban mobility applications. In [21], the authors even highlight how gravity models are sensible to the size of geographical clusters within the same city, in the case of dockless BSS, which may be problematic for interpretability purposes. With that in mind, we propose to assess the quality of a model by measuring how the fitted parameters vary as a function of scale within a given dataset, where the scale is determined by the number of geographical clusters. Using the exhaustive data available for the French city of Toulouse, the results demonstrate that the exponential Poisson model is by far the most consistent with respect to scale changes. Because the estimated values remain the most stable with varying pre-processing of the same dataset, we argue that robustness to scale and distance variations is a good indicator of greater parameter interpretability.

The exponential Poisson gravity model is then used to study the evolution of BSS use over time in Toulouse, as well as in the American cities of Washington and Chicago for which data are made public. A dynamic implementation is achieved by calibrating the model over a rolling five-week window, starting from January 2019 to December 2022. This approach reveals several common patterns on BSS use, such as distinct parameters values between work days and weekends, seasonal trends or adaptation to the lockdowns and restrictions during the Covid-19 crisis. The two main parameters of the model, denoted by  $\alpha$  and  $\beta$  throughout, can be naturally associated with two different aspects of the use of BSS, namely the nature of the trip in relation with the population density both at the origin and destination for  $\alpha$ , and the degree of reluctance to longer trips for  $\beta$ . Hence, analysing separately the time-series obtained for these two parameters reveals different types of patterns and provide a more flexible way to interpret the results. The same analysis was conducted on the cities of Los Angeles and Philadelphia, revealing similar behaviors at small time scales (e.g. comparing working days and weekends) but completely different long-term trends, which may be explained by the relatively small BSS infrastructure and differing policies in these two cities [17].

The paper is organized as follows. We discuss the general motivation and theoretical implications of the most common gravity models in Section 2. Using the Toulouse data, we show the extent to which the Poisson gravity model with an exponential relationship to distance is more accurate in terms of quality of fit and robustness to scale changes. This model is then tested on the American cities of Washington and Chicago, in addition to Toulouse, in Section 3. A dynamical analysis is conducted on these cities to study the evolution of the fitted parameters over the four-year period from January 2019 to December 2022. Some common patterns in the evolution of the model parameters provide some insight on the use of BSS in these cities. Limitations of the dynamical gravity model are shown by conducting the same analysis on the cities of Philadelphia and Los Angeles where different patterns emerge, some of which lacking a convincing explanation.

## 2 General overview of gravity models

We explore the different aspects of gravity models and illustrate their accuracy to describe BSS usage in urban areas. We discuss the two most common gravitation models variants which either assume a polynomial or exponential decay of the interaction between two nodes of an urban network in function of the distance. Then, we go further into the underlying assumptions of these models in terms of stochastic representation, which is directly linked to the inference method. To illustrate the different performances of the gravity models variants, we perform an analysis on the data for BSS usage in Toulouse over the course of 2019 on work days.

### 2.1 Regression function

Let  $X_{ij}, i, j = 1, \dots, N$  denotes a numerical vector measuring a level of interaction between nodes  $i$  and  $j$ . Inspired from Newton's law and introduced in [39], gravity models typically assume a relation of the form

$$X_{ij} \approx C \frac{(m_i m_j)^\alpha}{d_{ij}^\beta}, \quad (1)$$

where  $C, \alpha, \beta$  are unknown constants,  $m_i$  represents the *mass* at the node  $i$  and  $d_{ij}$  is the distance between  $i$  and  $j$  (in this paper, we only consider symmetrical cases where  $d_{ij}$  and  $d_{ji}$ ). For mobility data, the mass  $m_i$  is generally chosen as a socio-geographic variable strongly linked to the level of interaction such as the local density of population, level of wealth, etc... Gravity models are usually investigated from a graph theory perspective where the data are associated with a weighted graph  $G = (V, E)$  where  $V$  denotes the set of vertices or nodes (in general,  $V = \{1, \dots, N\}$ ) and  $E$  is the set of weighted edges. This point of view is especially relevant when the data are sparse, that is when several of the  $X_{ij}$ 's are equal to zero to indicate an absence of interaction. In this case, the set  $E$  only contains the pairs  $(i, j)$  for which  $X_{ij}$  is non-zero.

Many variants of the gravity models have been investigated in the literature, the most common example being to assume an exponential function of the distance [7]

$$X_{ij} \approx C(m_i m_j)^\alpha e^{-\beta d_{ij}}. \quad (2)$$

Another common variant is to treat the outward mass  $m_i^{\text{out}}$  and inward mass  $m_j^{\text{in}}$  as distinct variables, with potentially different exponents, investigated for instance in [7].

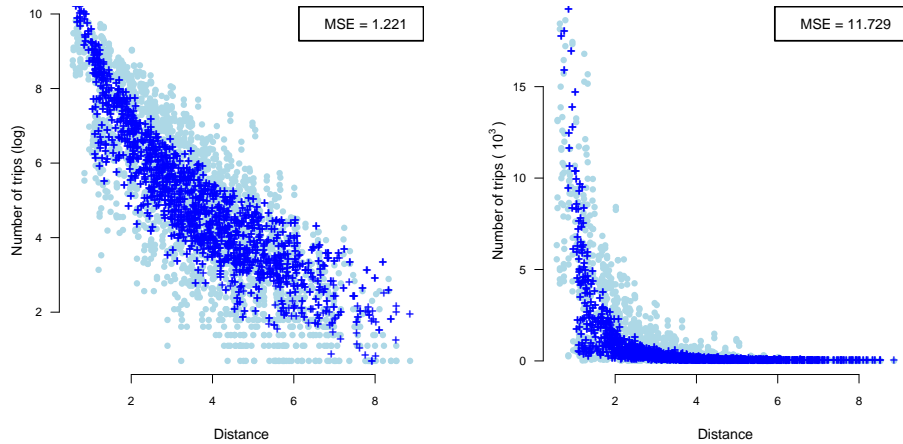


Figure 1: Comparison of the number of trips  $X_{ij}$  (light blue) and fitted values (dark blue) as a function of distance  $d_{ij}$  in the homoscedastic log-normal polynomial gravity model, with a y-axis in logarithmic scale (left) and regular (linear) scale (right). The mean square error (MSE) is calculated in both cases.

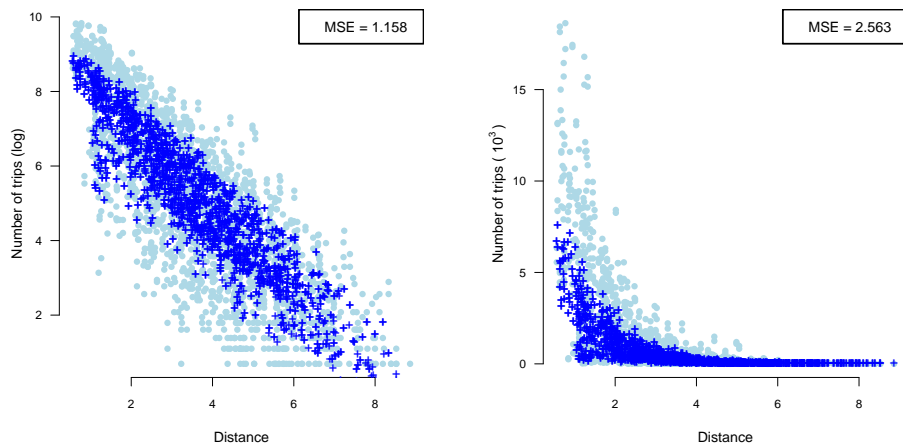


Figure 2: Comparison of the number of trips (light blue) and fitted values (dark blue) as a function of distance in the (homoscedastic) log-normal exponential gravity model, with a y-axis in logarithmic scale (left) and regular scale (right).

In Figures 1 and 2 (and in the following Figures 3 and 4), the light-blue points are always the same, representing the real values ( $\log(X_{ij})$  in the left plot and  $X_{ij}$  in the right plot). Visually, the logarithmic values appear to decrease linearly with respect to the distance  $d_{ij}$ , providing a more

faithful representation with the exponential relationship (Figure 1) compared to the polynomial one (Figure 2). Nevertheless, none of the log-normal models seem to provide a fully convincing fit (dark-blue points). Even with the seemingly more suited exponential relationship displayed in Figure 2, the model clearly underestimates the number of small-distance trips  $X_{ij}$  (right plot).

## 2.2 Conditional distribution

Parametric inference is generally achieved by linear regression on the logarithm of the data points  $X_{ij}$ . For example, in the *polynomial* gravity model (i.e. with the parameter  $\beta$  as an exponent of the distance), the linear regression approach corresponds to the relation

$$\log X_{ij} = \log C + \alpha \log(m_i m_j) - \beta \log(d_{ij}) + \epsilon_{ij}, \quad (i, j) \in E. \quad (3)$$

where the  $\epsilon_{ij}$ 's represent the error terms, generally treated as random variables with zero mean. Note that only the non-zero terms  $X_{ij}$  are conserved to avoid infinite values when taking the logarithm (hence the restriction to  $(i, j) \in E$ ). This means in particular that the absent interactions  $X_{ij} = 0$  are ignored for the inference.

The treatment of the error terms  $\epsilon_{ij}$ 's is an important, although sometimes neglected, aspect of the inference. Indeed, the criterion used to estimate the parameters is inextricably linked to the hypothesized distribution of the  $\epsilon_{ij}$ 's. For instance, the ordinary least square method, minimizing the quadratic criterion

$$R(C, \alpha, \beta) = \sum_{(i,j) \in E} \left( \log X_{ij} - (\log C + \alpha \log(m_i m_j) - \beta \log d_{ij}) \right)^2, \quad (4)$$

is generally used as the default inference method, ignoring that the criterion should be chosen depending on the distribution of the error (typically, least square criteria are unreliable if the distribution of the errors is heavy tailed). From a purely probabilistic point of view, this criterion is equivalent to the maximum likelihood when  $X_{ij}$  has a log-normal distribution conditionally to the masses  $m_i, m_j$  and distance  $d_{ij}$ , which we denote by

$$X_{ij} \sim LN(\Phi(m_i, m_j, d_{ij}), \sigma^2(m_i, m_j, d_{ij})).$$

The first parameter  $\Phi(m_i, m_j, d_{ij})$  is the expectation of  $\log(X_{ij})$  conditionally to  $m_i, m_j, d_{ij}$  and thus depends on the type of gravity model used :

$$\Phi(m_i, m_j, d_{ij}) = \begin{cases} \log C + \alpha \log(m_i m_j) - \beta \log(d_{ij}) & \text{(polynomial)} \\ \log C + \alpha \log(m_i m_j) - \beta d_{ij} & \text{(exponential)} \end{cases}$$

The variance parameter  $\sigma^2(m_i, m_j, d_{ij})$  depends on the amount of detail one is willing to invest in the model and inference method. The most simple case where all variances are assumed to be equal (homoscedasticity assumption) justifies the use of the ordinary least square criterion of Equation (4) and generally provides consistent estimators of the parameters under reasonable conditions. It is the most popular approach due to its simplicity and efficiency. More intricate criteria such as weighted least squares are justified in the log-normal scenario if one has some prior knowledge on the variances parameters and how they relate to the explanatory variables  $m_i, m_j, d_{ij}$ , or to further improve the inference method to account for the heteroscedasticity.

Generalized linear models (GLM) are alternatives to Gaussian linear regression associated to other conditional distributions for the variable of interest  $X_{ij}$ . A widely used model for integer

response is the so-called Poisson regression model where  $X_{ij}$  is assumed to have a Poisson distribution (denoted by  $\mathcal{P}(\cdot)$ ) conditionally to the distance  $d_{ij}$  and masses  $m_i, m_j$ . When coupled with a gravity model that provides a parametric expression of the conditional expectation function  $\Phi(m_i, m_j, d_{ij}) = \mathbb{E}(X_{ij} | m_i, m_j, d_{ij})$ , the Poisson regression model corresponds to the following stochastic representation of the conditional distribution of  $X_{ij}$  :

$$X_{ij} \sim \mathcal{P}(\Phi(m_i, m_j, d_{ij})).$$

In the Poisson model, the parameter  $\Phi(m_i, m_j, d_{ij})$  dictates both the behavior of the expectation and variance, making this setting somewhat less flexible than the log-normal distribution where an additional variance parameter is present. Nevertheless, it does provide a better fit for the BSS data.

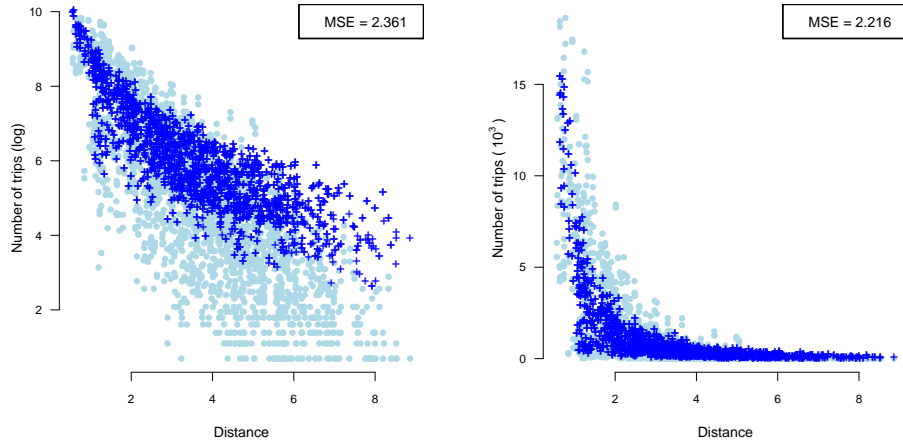


Figure 3: Comparison of the number of trips (light blue) and fitted values (dark blue) as a function of distance in the Poisson polynomial gravity model, with a y-axis in logarithmic scale (left) and regular scale (right).

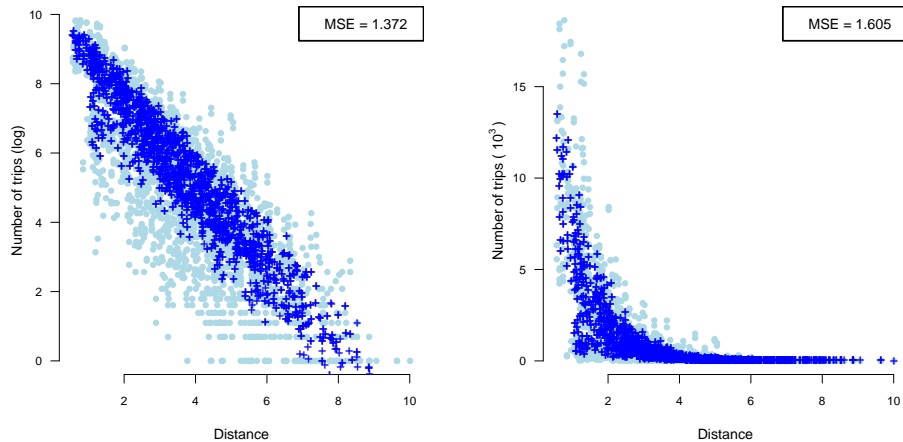


Figure 4: Comparison of the number of trips (light blue) and fitted values (dark blue) as a function of distance in the Poisson exponential gravity model, with a y-axis in logarithmic scale (left) and regular scale (right).

Unsurprisingly, the polynomial relationship to distance remains ill-suited for the logarithmic values of the number of trips, as visible on the left plot in Figure 3. The Poisson exponential model displayed in Figure 4 does however appear to satisfactorily represent the logarithmic and regular values of the number of trips  $X_{ij}$ .

Gravity model	Polynomial		Exponential	
	log-normal	Poisson	log-normal	Poisson
MSE (log)	1.221	2.361	<b>1.158</b>	1.372
MSE ( $\times 10^3$ )	11.729	2.216	2.563	<b>1.605</b>

Table 1: Mean square errors of the various gravity models.

The log-normal assumption on the conditional distribution corresponds to a minimization of the MSE on the logarithmically transformed data, which explains why the minimal MSE is achieved for these models. Nevertheless, these produce a bad fit on the raw (non-logarithmic) data, especially in the polynomial gravity model. Conversely, Poisson models are more inclined to minimize the MSE on the untransformed data, although it is not the criterion used to calibrate the parameters. The exponential gravity model appears to be much more suited to the BSS data for both distributions, with the Poisson model providing the overall best results.

### 2.3 Robustness to scale change

We now discuss the relevance of a gravity model for mobility via its robustness (i.e. stability) to scale change in the data processing. We argue that to be able to interpret the values of estimated parameters in a given model in term of characteristics and features in the dataset, these values should hardly be impacted by a pre-processing of the data. Typically, the  $\beta$  parameter of the gravity models in Equations (1) and (2) is linked to the distances of the BSS trips. Its estimated value can be



interpreted as a tolerance coefficient for the users to do long distance trips: a higher value conveys a higher tendency to travel longer distances. If the model is well defined, we expect this value  $\beta$  to be impervious to scale variations in a preliminary clustering of the BSS stations, provided these scale variations are made for visualization and inference purposes only, with no intention of transforming the data. This unwanted phenomenon is visible in [21] where the estimated parameters of the gravity model are strongly dependent on the scale of the space partitioning, making the interpretation of these values relevant relative to each other within each model but not in an absolute sense.

We use this robustness criterion to measure and compare the quality of interpretability of the parameters in the four gravity models stemming from exponential or polynomial functions of the distances and log-normal or Poisson distributions, discussed in the previous section. The scale change arises from a preliminary clustering of the stations based on geographical proximity. While BSS stations could be dealt with individually in the model, it is only natural to gather the stations into small clusters since nearby stations are usually interchangeable from the user perspective and are expected to play the same role in the scheme of things. From the practitioner point of view, regrouping the stations into small clusters is a way to reduce the data variability at each location while reducing the size of the network for an easier implementation of the model. Provided that the clusters are built based on geographical proximity and are reasonably distributed in the network, variations of the clusters average size between say 3 and 10 stations should not cause a transformation in the data and hence, should not modify considerably the model's parameters.

We see in Figure 5 that the Poisson exponential model is extremely robust to scale change for the parameter  $\beta$ , followed by the log-normal exponential model. The polynomial relationship in the distance induces a strong dependency to scale, as can be expected from the behaviour of the regression function for near-zero distances. About the  $\alpha$  parameter, the variations as a function of scale are lower with, here again, a worse performance for the polynomial relationship. Finally, the Poisson exponential model has the better quality of fit among the four models, regardless of scale.

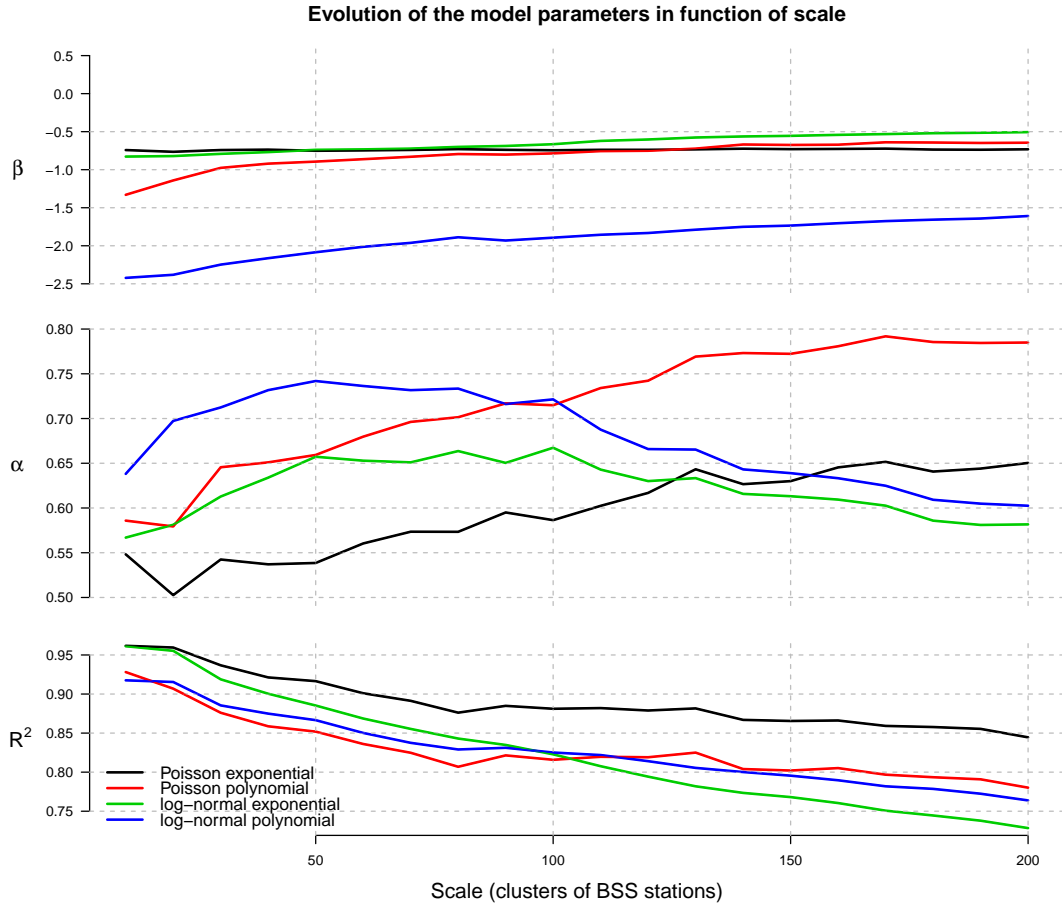


Figure 5: Evolution of the models parameters as a function of the scale.

In conclusion, the exponential Poisson model is by far the most robust to scale change among the four considered models. It also provides the highest quality of fit for the data when aiming for a compromise between MSE and logarithmic MSE (see Table 1) as well as for the pseudo  $R^2$  criterion.

The three parameters of the model can be associated to distinct characteristics, which makes it particularly suited for interpretability. Precisely :

- The  $\alpha$  parameter is associated to the density of population over each station (or cluster of stations). It is linked to the differences of population density between locations for high frequency trips. A higher prevalence of trips from low-density locations is expected to induce a larger value of  $\alpha$ , all other things being equal.
- The  $\beta$  parameter is associated to the distances of the trips. A high value of  $\beta$  can be interpreted as a higher prevalence of long-distance trips and can thus be associated with a degree of the users reluctance to longer trips.

- The  $C$  parameter (the intercept) is associated to the volume of BSS trips over the whole considered area. It is expected to fluctuate between high usage periods (e.g. spring) and low usage periods (e.g. winter vacations), other things being equal. In practice however, this parameter is highly sensitive to small changes in distances and population density which makes it difficult to interpret independently from the other two parameters.

### 3 Comparative analysis and dynamical modelling

We tested the model on the French city of Toulouse and the American cities of Chicago and Washington, for which we have each BSS trip performed between 2019 and 2022. Toulouse data are provided by the BSS company JC Decaux, and Chicago and Washington data are in free and open access []. The data, which contains all trips made with start-end times and origin-destination bike stations, was processed as follows.

- Only the trips lasting between 2 minutes and 1 hour are included as a simple and convenient way to eliminate outliers.
- BSS stations were gathered geographically into small clusters in each city (100 clusters for Toulouse, 250 for Chicago and 200 for Washington). The number of clusters was chosen so as to have an average of around 3-4 stations per cluster in each city, in order to reduce variability in the data. Due to the lack of bicycles availability that can be observed in some stations at rush hours (i.e. the station is either full or empty which forces the user to move to the nearest station), nearby stations often play the same role in the network which further justifies the clustering process. Note that, as discussed in the previous section, slight variations of the number of clusters do not impact significantly the estimated parameter values. The trips from one station to another within each cluster are omitted due to the impossibility to accurately account for the distance within a cluster.
- An approximation of the local population density based on BSS activity was used for the variable  $m_i$ , defined as the total number of trips starting or arriving at cluster  $i$  during the reference year 2019. This value is used for the whole period 2019-2022 in the three cities.

Based on the preliminary analysis of Section 2, we adjust the data using the exponential Poisson model where the number of trips  $X_{ij}$  between stations clusters  $i$  and  $j$  is assumed to follow a Poisson distribution with parameter

$$\Phi(m_i, m_j, d_{ij}) = C(m_i m_j)^\alpha e^{-\beta d_{ij}}.$$

The quality of fit of the model is measured by the (pseudo)  $R^2$  defined as

$$R^2 = 1 - \frac{D}{D_0}$$

where  $D$  and  $D_0$  are respectively the deviance of the Poisson model and the null model (see for instance [9] for the technical details).

While it is expected to vary with the total number of trips, the parameter  $C$  is highly sensible to fluctuations of the other parameters  $\alpha, \beta$ . Its main purpose is thus to optimize the model fit and lacks an interesting interpretation. For this reason, the estimated values for  $C$  are omitted from the following analysis. An overview of the different behavior in BSS usage in these three cities is given in Table 2 below, which shows the estimated values for the parameters  $\alpha, \beta$  for the years 2019 to 2022.

		2019	2020	2021	2022
Toulouse	$\alpha$	0.568	0.578	0.585	0.580
	$\beta$	0.786	0.745	0.739	0.757
	$R^2$	0.870	0.868	0.876	0.873
Washington	$\alpha$	0.451	0.503	0.559	0.465
	$\beta$	0.763	0.609	0.631	0.704
	$R^2$	0.875	0.815	0.853	0.889
Chicago	$\alpha$	0.583	0.584	0.785	0.942
	$\beta$	0.574	0.467	0.311	0.251
	$R^2$	0.839	0.818	0.753	0.736

Table 2: Estimated parameters and pseudo  $R^2$  in the three cities from year 2019 to 2022. All are non-zero with high significance (p-value  $< 10^{-3}$ ).

In Toulouse, the  $\alpha$  parameter remains stable during the four-year period. Since  $\alpha$  is associated to the population density at a trip endpoints, this could suggest that the type of BSS trips has not changed in this period. The fact that there were few modifications in the BSS infrastructure in Toulouse between 2019 and 2022 should also have contributed to the stability of this parameter. On the contrary, the  $\beta$  parameter has notably decreased during the pandemic before going back up in 2022 while remaining below its pre-pandemic value. This can be explained by a change in the BSS users behavior in 2020 and 2021, where the lower reluctance to longer trips can be caused for instance by the will to avoid crowd proximity in public transportation. We observe a trend back towards the pre-pandemic habits in 2022, confirming the findings of [13], although the change in behavior clearly continued even after all mobility restrictions were lifted. This phenomenon becomes even more evident in the temporal analysis discussed further. The model  $R^2$  remains high, around 0.87, and stable during the four years, despite the change in parameter values.

In Washington, the effect of the pandemic is visible for parameters  $\alpha$  and  $\beta$ , indicating both a change in the type of trips and users habits. The high values of  $\alpha$  during the pandemic years could be interpreted as a result of the policy restrictions, particularly the ban on public transportation, leading to a sudden interest from individuals other than regular BSS-users. Thus, the BSS usage during the pandemic undergoes a notable change as visible through the difference in the estimated values of both  $\alpha$  and  $\beta$  during the years 2020 and 2021. By 2022,  $\alpha$  has returned to its 2019 value while  $\beta$  remains significantly below.

In Chicago, the impact of the pandemic is not clearly revealed by the evolution of the two parameters. The increase of  $\alpha$  after the pandemic is attributed to the rise in BSS usage, from 3.3 million to 5.5 million trips, and the expansion of the BSS network from 581 stations in 2019, to 647 in 2021, and to more than 700 in 2022 [17]. The extended network and higher usage may also explained the slight decline in fit quality going from a pseudo  $R^2$  of around 0.87 in 2019 to 0.7 in 2022. What is particularly notable in this city is the continuing decrease in the values of  $\beta$ , indicating users' willingness to undertake long-distance trips with shared-bikes, likely due to the gradual introduction of electric bikes.

In Table 3, we compare the results obtained for working days and weekends separately, as well as for two distinct three-month periods, namely winter (January, February, March) and summer (June, July, August) in the benchmark year 2019.

		Winter 2019			Summer 2019		
		$\alpha$	$\beta$	$R^2$	$\alpha$	$\beta$	$R^2$
Toulouse	Working day	0.550	0.768	0.818	0.592	0.664	0.835
	Rest day	0.531	0.714	0.785	0.565	0.617	0.790
Washington	Working day	0.461	0.794	0.887	0.467	0.773	0.896
	Rest day	0.475	0.794	0.889	0.482	0.757	0.886
Chicago	Working day	0.597	0.589	0.740	0.632	0.517	0.796
	Rest day	0.262	0.541	0.784	0.489	0.490	0.769

Table 3: Comparison of the estimated parameters in summer vs. winter and working day vs. rest day in 2019.

The changes measured between the different time periods are compatible with the previous interpretations in Toulouse and Chicago. The higher percentage of peripheral trips, longer and connecting less densely populated locations, is consistent with the smaller values of the two parameters  $\alpha, \beta$  on the weekends. The gap for  $\alpha$  is particularly visible during the harsh Chicago winter with the smallest measured value on rest days. The differences between seasons are also high in these two cities. Moreover, long-distance trips are more frequent in summer, as conveyed by the smaller  $\beta$  parameter. On the other hand, we observe very few variations in Washington, neither between working days and rest days nor between summer and winter.

### 3.1 Temporal evolution

Studying the temporal evolution of travelers' behaviors in bike-sharing systems (BSS) provides valuable insights into urban mobility dynamics and underlying societal trends. In this purpose, we explore the evolution of the two main parameters using BSS data for Toulouse, Washington and Chicago from 2019 to 2022. The values are obtained by adjusting the model on the data restricted to a rolling five-week window, analysing working days and weekends separately, over the four-year period. This study allows us to observe fluctuations in BSS usage, as well as users' responses to major events such as the COVID-19 pandemic and changes in BSS infrastructure. By examining annual variations in key parameters such as  $\alpha$  and  $\beta$ , we can identify emerging trends, seasonal patterns, and the impacts of policy interventions on the travel habits of urban residents.

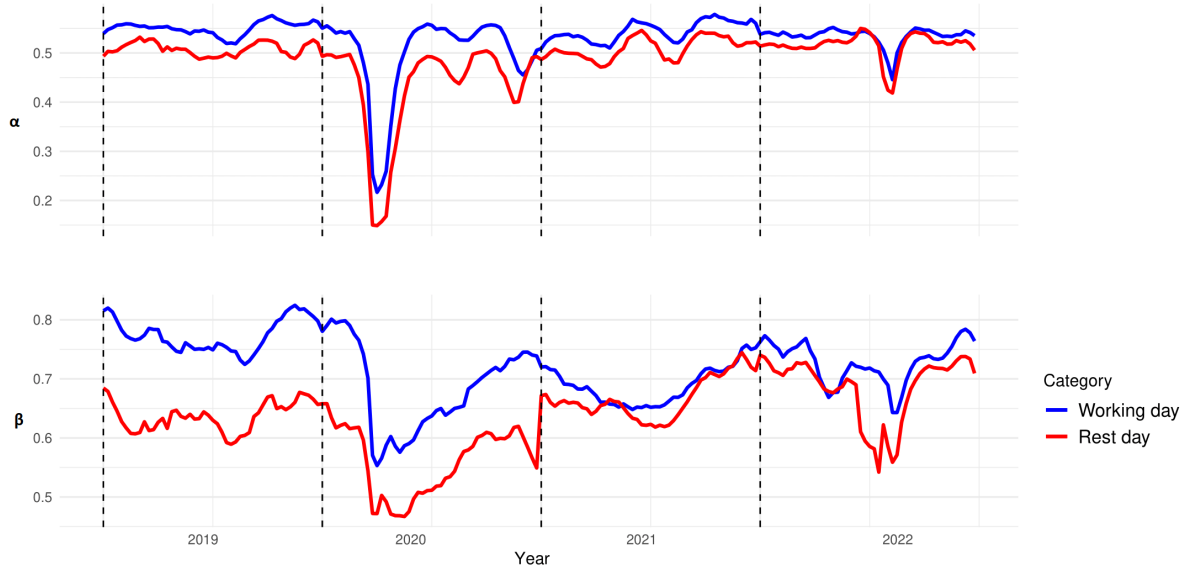


Figure 6: Evolution of the estimated  $\alpha$  and  $\beta$  coefficient as a function of the week over the 4-year period in Toulouse city.

As we can see in Figure 6, both parameters are systematically higher during the 5 first days of the week, confirming the difference of BSS usage between a working day and a weekend day. The higher gap for  $\beta$  between working days and weekends in 2019 seems to narrow down from 2021 onwards. The steep decline in the parameter values, caused by the mobility restrictions imposed by the french government at the start of the Covid-19 pandemic, is clearly visible in the second semester of 2020.

The changes in values for  $\alpha$  appears to be directly caused by mobility government restrictions, with a rapid return to a normal level once these restrictions are lifted. This phenomenon is particularly visible during the two lock-downs of 2020 (from March 17th to May 3rd and from October 30th to December 15th). On the other hand, the variations of the  $\beta$  parameter show some inertia for the period just following each lock-down, suggesting a more long-term change of mindset in the BSS users, although they seem to have eventually returned to a normal level in 2022 <sup>1</sup>.

<sup>1</sup>Investigation of the somewhat strange behavior observed in the summer of 2022 led us to believe that it was caused by an anomaly when collecting the data. For this reason, it is not discussed in the analysis.

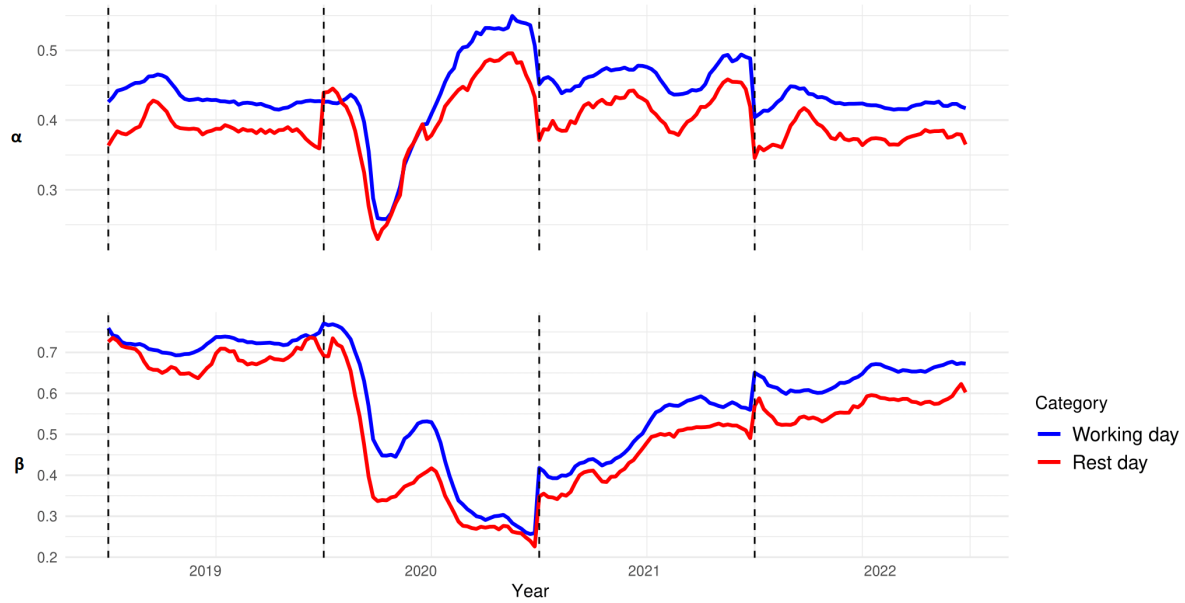


Figure 7: Evolution of the estimated  $\alpha$  and  $\beta$  coefficient in function of the week of Washington.

The conclusions are similar in Washington, as can be seen in Figure 7. The values of  $\alpha$  and  $\beta$  remain greater during working days, regardless of the time of year over the whole period. The higher inertia for  $\beta$ , which slowly increases after its initial drop in 2020, is also present here while the variations for  $\alpha$  are more rapid. The evolution of the two parameters during the pandemic period is however noticeably different from Toulouse, e.g. at the end of 2020, probably due to the different timing of the restrictions in this city.

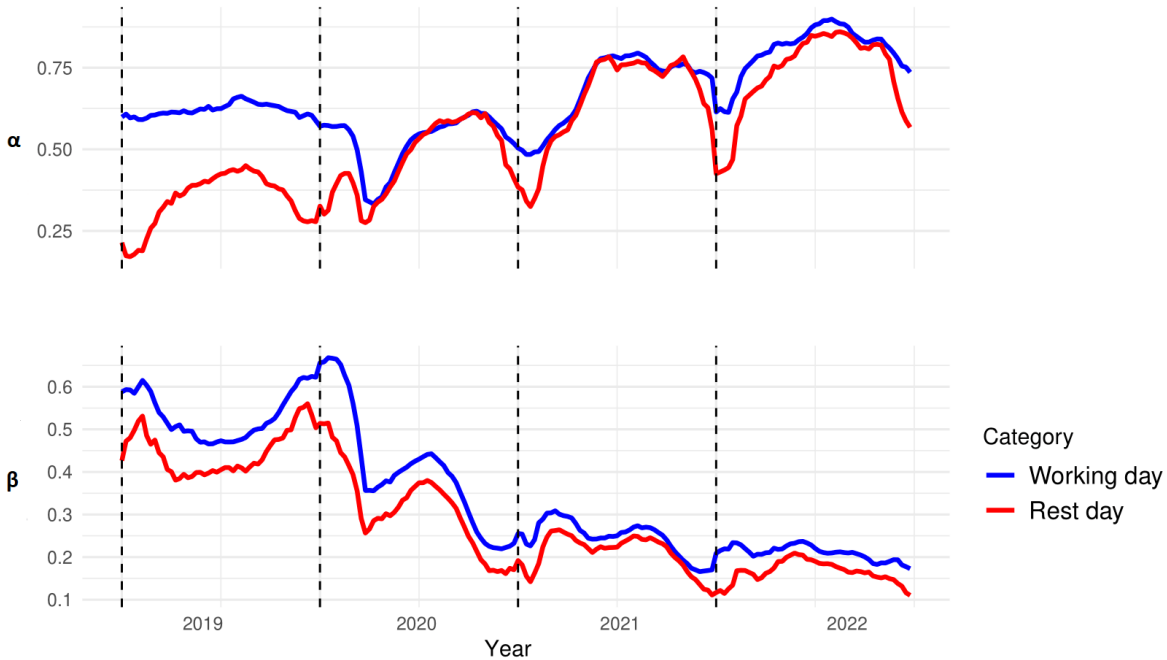


Figure 8: Evolution of the estimated  $\alpha$  and  $\beta$  coefficient in function of the week of Chicago.

In Chicago, a clear difference in the evolution of  $\alpha$  values in 2019 is noticeable between working and rest days, but for the following three years,  $\alpha$  values are mostly higher during working days. The  $\alpha$  increasing in Chicago is attributed to the increase in BSS usage (from 3.3 million to 5.5 million). Here again, the estimated values for  $\beta$  are smaller during rest days, and there is a continuous decrease during and after the pandemic, indicating users' lack of concern about making long distances with BSS. Unlike the two previously mentioned cities, we are no longer in the pattern of returning to pre-pandemic behavior; habits during the pandemic remain after the restriction. This is most likely attributed to incentive and expansion programs of the BSS system in Chicago, notably the introduction of electric bicycles. While a seasonal effect is visible for both parameters, the parameter  $\alpha$  is by far the more affected by the harsh weather conditions in Chicago, especially during the weekends.

### 3.2 Model limitation : the cases of Philadelphia and Los Angeles

The same analysis was conducted on the cities of Philadelphia and Los Angeles with mixed results. Urban cycling mobility is less developed in these two cities, with approximately 0.2 million BSS trips annually, a number smaller by an order of magnitude compared to the three previous cases.



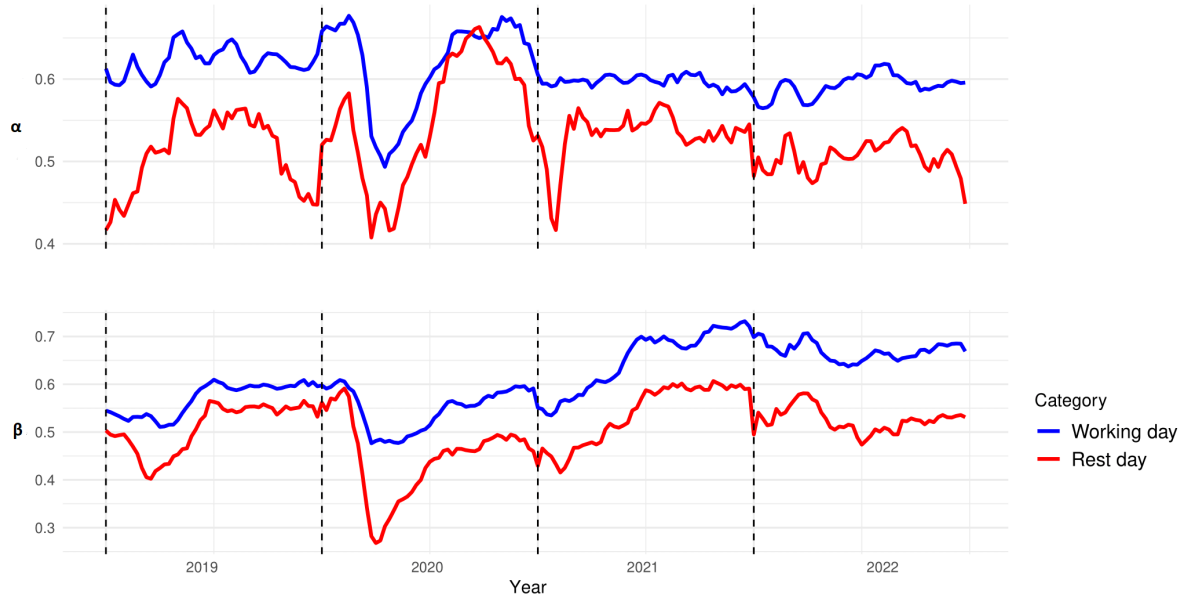


Figure 9: Evolution of the estimated  $\alpha$  and  $\beta$  coefficient in function of the week of Philadelphia.

We can see in Figure 9 that the pandemic and mobility restrictions have affected the BSS use in Philadelphia, as illustrated for instance by the change point visible in the second trimester of 2020, although the impact seems to be less pronounced. Despite the somewhat more chaotic behavior of the  $\alpha$  time series, most probably due to the smaller sample size causing a higher variability of the estimated values, a fast return to normal seems to occur following the end of the restrictions. On the contrary, the  $\beta$  coefficient exhibits a slower return to stability, with the 2022 values remaining higher than 2019 for working days in 2022 and comparable on the weekends.

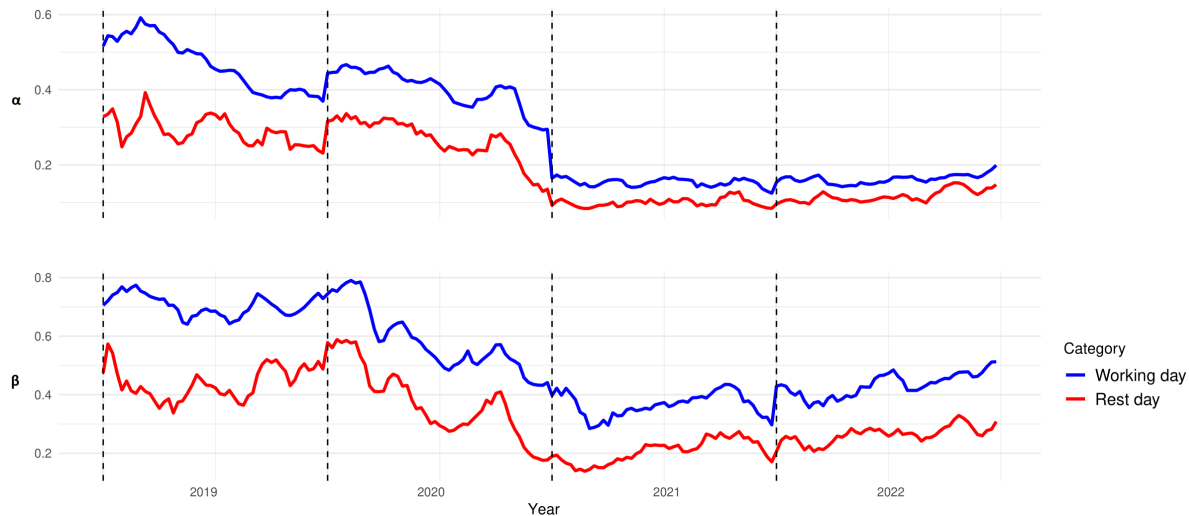


Figure 10: Evolution of the estimated  $\alpha$  and  $\beta$  coefficient in function of the week of Los Angeles.

In Los Angeles (Figure 10), the evolution of the calibrated parameters  $\alpha, \beta$  highly differs from the previous cases, the only similar phenomenon being the higher values of the two coefficients on working days. The impact of the Covid-19 pandemic is less pronounced and no brutal change occurs during the four-year period. We are unable to provide a clear explanation to the general trend characterized in all four time series by a gradual decay during the year 2020 followed by a hardly noticeable increase the two following years. Note that the city of Los Angeles is one of the largest cities in the world in term of surface area and its BSS infrastructure is composed of two large groups of stations several kilometers apart, which may cause such differences with the other cities studied.

## 4 Concluding remarks

Gravity models accurately describe the use of bike sharing systems in urban areas when accounting for the local density of population and distances between stations. Among the main variants, the Poisson distributed gravity model with exponential relationship with respect to distance appears as the most suited to describe BSS usage, based on quality of fit and robustness to scale change. The two main coefficients  $\alpha, \beta$  of the model are associated respectively to the type of trips (e.g. work related, from the city center to the periphery etc...) through the population density of the origin and destination, and to the BSS users' habits through the traveled distance.

When calibrated over a rolling time window, the models allow to describe the evolution of BSS usage over time and exhibit particular behaviors depending on the day of the week, the season or even exceptional scenarios such as the Covid-19 pandemic. Because gravity models are easily interpretable, the estimated values of the coefficients  $\alpha$  and  $\beta$  over the different time periods provide some insight on the evolution of BSS usage.

A study was conducted on the french city of Toulouse and the American cities of Chicago, Wash-

ington, Philadelphia and Los Angeles, highlighting several common patterns. Most notably, the values of  $\alpha$  and  $\beta$  are always higher the weekends regardless of the time of the year, which is consistent with the higher percentage of peripheral trips and a more casual use on rest days. With the exception of Los Angeles, the mobility restrictions imposed in the first half of 2020 caused a sudden change in the BSS use, which took time to return to its pre-pandemic standard. The effect of the restrictions are noticeably different for the two coefficients, with  $\beta$  taking more time to return to a stability, even long past the end of the restrictions. This suggests that the pandemic had a durable effect on the population mindset regarding the use of BSS, and specifically on the willingness to travel longer distances by bike. In most cases, this change seems to have vanished almost completely by the end of 2022. In the case of Chicago, the gradual introduction of electrical bikes over the 2020-2022 period induces a continuous decrease in the  $\beta$  coefficient, consistent with an increasing tendency by the BSS users to travel longer distances.

## References

- [1] Diego Maria Barbieri, Baowen Lou, Marco Passavanti, Cang Hui, Inge Hoff, Daniela Antunes Lessa, Gaurav Sikka, Kevin Chang, Akshay Gupta, Kevin Fang, et al. Impact of covid-19 pandemic on mobility in ten countries and associated perceived risk for all transport modes. *PloS one*, 16(2):e0245886, 2021.
- [2] Julien Bert, Daniel Schellong, Markus Hagenmaier, David Hornstein, Augustin K Wegscheider, Thomas Palme, et al. How covid-19 will shape urban mobility. <https://www.bcg.com/publications/2020/how-covid-19-will-shape-urban-mobility>, 2020.
- [3] Richard J Bouchard and Clyde E Pyers. Use of gravity model for describing urban travel. *Highway Research Record*, 88(196):5, 1965.
- [4] Ralph Buehler and John Pucher. Covid-19 and cycling: a review of the literature on changes in cycling levels and government policies from 2019 to 2022. *Transport Reviews*, 44:299–344, 04 2023.
- [5] Xinwei Chai, Xian Guo, Jihua Xiao, and Jie Jiang. Analysis of spatial-temporal behavior pattern of the share bike usage during covid-19 pandemic in beijing. *arXiv preprint arXiv:2004.12340*, 2020.
- [6] Ho Fai Chan, Ahmed Skali, David A Savage, David Stadelmann, and Benno Torgler. Risk attitudes and human mobility during the covid-19 pandemic. *Scientific reports*, 10(1):1–13, 2020.
- [7] Yanguang Chen. The distance-decay function of geographical gravity model: Power law or exponential law? *Chaos, Solitons & Fractals*, 77:174–189, 2015.
- [8] Paul DeMaio. Bike-sharing: History, impacts, models of provision, and future. *Journal of public transportation*, 12(4):3, 2009.
- [9] Annette J Dobson and Adrian G Barnett. *An introduction to generalized linear models*. Chapman and Hall/CRC, 2018.
- [10] Ezgi Eren and Volkan Emre Uz. A review on bike-sharing: The factors affecting bike-sharing demand. *Sustainable Cities and Society*, 54:101882, 2020.
- [11] Elliot Fishman. *Bike Share*. Routledge, 2019.
- [12] Robin Flowerdew and Murray Aitkin. A method of fitting the gravity model based on the poisson distribution. *Journal of Regional Science*, 22(2):191–202, 1982.

- [13] Angelo Furno, Bertrand Jouve, Bruno Revelli, and Paul Rochet. Impact of the COVID-19 pandemic on bike-sharing uses in two French towns: the cases of Lyon and Toulouse. working paper or preprint, December 2023.
- [14] Segun Goh, Keumsook Lee, Jong Soo Park, and M. Y. Choi. Modification of the gravity model and application to the metropolitan seoul subway system. *Phys. Rev. E*, 86:026102, 2012.
- [15] Eva Heinen, Bert Van Wee, and Kees Maat. Commuting by bicycle: an overview of the literature. *Transport reviews*, 30(1):59–96, 2010.
- [16] Songhua Hu, Chenfeng Xiong, Zhanqin Liu, and Lei Zhang. Examining spatiotemporal changing patterns of bike-sharing usage during covid-19 pandemic. *Journal of Transport Geography*, 91:102997, 2021.
- [17] Zhaoyu Kou and Hua Cai. Comparing the performance of different types of bike share systems. *Transportation research part D: transport and environment*, 94:102823, 2021.
- [18] Oh-Hyun Kwon, Inho Hong, Woo-Sung Jung, and Hang-Hyun Jo. Multiple gravity laws for human mobility within cities. 2023.
- [19] Aoyong Li, Pengxiang Zhao, Haitao He, and Kay W Axhausen. Understanding the variations of micro-mobility behavior before and during covid-19 pandemic period. *Arbeitsberichte Verkehrs- und Raumplanung*, 1547, 2020.
- [20] Haojie Li, Yingheng Zhang, Manman Zhu, and Gang Ren. Impacts of covid-19 on the usage of public bicycle share in london. *Transportation Research Part A: Policy and Practice*, 150:140–155, 2021.
- [21] Ruiqi Li, Shuai Gao, Ankang Luo, Qing Yao, Bingsheng Chen, Fan Shang, Rui Jiang, and H Eugene Stanley. Gravity model in dockless bike-sharing systems within cities. *Physical Review E*, 103(1):012312, 2021.
- [22] Cyrille Médard De Chardon. *Bicycle Sharing*, volume 6, pages 19–30. Elsevier, 1 edition, May 2021.
- [23] Francesco Pase, Federico Chiariotti, Andrea Zanella, and Michele Zorzi. Bike sharing and urban mobility in a post-pandemic world. *Ieee Access*, 8:187291–187306, 2020.
- [24] John Pucher, Jennifer Dill, and Susan Handy. Infrastructure, programs, and policies to increase bicycling: an international review. *Preventive medicine*, 50:106–125, 2010.
- [25] Susan Shaheen, Stacey Guzman, and Hua Zhang. Bikesharing across the globe. *City cycling*, 183, 2012.
- [26] Wen-Long Shang, Jinyu Chen, HuiBo Bi, Yi Sui, Yanyan Chen, and Haitao Yu. Impacts of covid-19 pandemic on user behaviors and environmental benefits of bike sharing: A big-data analysis. *Applied Energy*, 285:116429, 2021.
- [27] Filippo Simini, Gianni Barlacchi, Massimiliano Luca, and Luca Pappalardo. A deep gravity model for mobility flows generation. *Nature Communications*, 12, 11 2021.
- [28] Jie Song, Liye Zhang, Zheng Qin, and Muhamad Azfar Ramli. Where are public bikes? the decline of dockless bike-sharing supply in singapore and its resulting impact on ridership activities. *Transportation Research Part A: Policy and Practice*, 146:72–90, 2021.
- [29] Jie Song, Liye Zhang, Zheng Qin, and Muhamad Azfar Ramli. Spatiotemporal evolving patterns of bike-share mobility networks and their associations with land-use conditions before and after the covid-19 outbreak. *Physica A: Statistical Mechanics and its Applications*, 592:126819, 2022.
- [30] Hyungun Sung. Causal impacts of the coronavirus-19 pandemic on daily ridership of public bicycle sharing in seoul. *SSRN Electronic Journal*, 89:104344, 2022.

- [31] João Filipe Teixeira and Miguel Lopes. The link between bike sharing and subway use during the covid-19 pandemic: The case-study of new york’s citi bike. *Transportation research interdisciplinary perspectives*, 6:100166, 2020.
- [32] João Filipe Teixeira, Cecília Silva, and Frederico Moura e Sá. The motivations for using bike sharing during the covid-19 pandemic: Insights from lisbon. *Transportation research part F: traffic psychology and behaviour*, 82:378–399, 2021.
- [33] João Filipe Teixeira, Cecília Silva, and Frederico Moura e Sá. Empirical evidence on the impacts of bikesharing: a literature review. *Transport reviews*, 41(3):329–351, 2021.
- [34] João Teixeira, Cecília Silva, and Frederico Moura e Sá. Potential of bike sharing during disruptive public health crises: A review of covid-19 impacts. *Transportation Research Record Journal of the Transportation Research Board*, pages 1–13, 2023.
- [35] CA Thompson, K Saxberg, J Lega, D Tong, and HE Brown. A cumulative gravity model for inter-urban spatial interaction at different scales. *Journal of Transport Geography*, 79:102461, 2019.
- [36] Haoyun Wang and Robert B Noland. Bikeshare and subway ridership changes during the covid-19 pandemic in new york city. *Transport policy*, 106:262–270, 2021.
- [37] Shuo Zhang, Li Chen, and Yingzi Li. Shared bicycle distribution connected to subway line considering citizens’ morning peak social characteristics for urban low-carbon development. *Sustainability*, 13(16):9263, 2021.
- [38] Pengjun Zhao, Haoyu Hu, Liangen Zeng, Jun Chen, and Xinyue Ye. Revisiting the gravity laws of inter-city mobility in megacity regions. *Science China Earth Sciences*, 66(2):271–281, 2023.
- [39] George Kingsley Zipf. The p1 p2 d hypothesis: On the intercity movement of persons. *American Sociological Review*, 11:677, 1946.