**Submitted to *Manufacturing & Service Operations Management*
manuscript**

The Impact of Congestion and its Antecedents: Evidence from Reviews

Alejandro Serrano
IESE Business School, University of Navarra, Av. Pearson 21, 08034 Barcelona, Spain. aserrano@iese.edu

F. Javier López
Dpto. Métodos Estadísticos and BIFI, Universidad de Zaragoza, Spain. javier.lopez@unizar.es

Víctor Martínez-de-Albéniz
IESE Business School, University of Navarra, Av. Pearson 21, 08034 Barcelona, Spain. valbeniz@iese.edu

Problem definition: In service management, congestion is known to decrease user satisfaction, everything else equal. Theory suggests that sensitivity to waiting is driven by the service configuration, yet there is little empirical evidence about its antecedents. **Methodology**: We use public reviews to measure the incidence of congestion in hundreds of touristic services, and to infer the impact that congestion has on satisfaction. **Results**: We show that the nature of the service is the most important driver of sensitivity to waiting, being highest in high-price, entertainment sites, such as amusement parks and zoos, whereas it is smallest in free, cultural sites, such as historical locations. Interestingly, sensitivity to waiting is reduced when the provider delivers excellent service on intangibles such as staff attentiveness or facilities cleanness. Furthermore, sensitivity is higher for locals in comparison with tourists. **Managerial implications**: Our results thus suggest that it is possible to offer high-satisfaction service experiences with high waiting, provided that other service outcomes are well executed.

Key words: service operations, text mining, service quality, congestion, touristic attractions.
History: Submitted: November 24, 2023.

1. Introduction

Since the seminal works of Maister (1985) and Larson (1987), it is well known that service satisfaction is eroded by the congestion experienced by a consumer. The theory suggests that waiting is disliked because it creates anxiety of not knowing when the service will take place, and imposes the cost of wasting one's precious time in a no-value-added queue. As a result, service managers should strive to inform about expected wait, turn wait into occupied in-service time and make queues socially just. These predictions have been supported by experimental studies (Rafaeli et al. 2002, Zhou and Soman 2008, Voorhees et al. 2009).

Unfortunately, existing studies usually provide evidence of the impact of congestion on service outcomes in unchanged environments, in which the service configuration has not been affected during the experiment, and only the amount of congestion has been altered. This implies that extant research cannot inform service managers about better service design alternatives, so as to diminish their exposure to congestion. To do this, an ideal experiment would compare various designs (say, A and B) and confront satisfaction obtained under low and high congestion levels. A more robust design would be such that, at low levels of congestion, A and B would deliver similar satisfaction levels, but design A would deliver higher satisfaction compared to B at high levels of congestion. A should be preferable to B because it would allow the service provider to handle higher throughput without sacrificing satisfaction. Setting up such an experiment would be ideal to learn how to design the service, but it is rather complex compared to standard lab experiments. In particular, one problematic question would be that of external validity and generalization to real service contexts. Another experimental option would be to run A/B tests in actual settings in various service concepts, such as bank branches or restaurants; but this requires the participation of a large organization, which, to the best of our knowledge, has not been possible to date.

In this paper, we propose an alternative approach, which relies on the use of retrospective data across a large number of service locations. The diversity of service design factors across these locations provides variation along the dimensions of theoretical interest, e.g., whether well-designed and well-run services make consumers penalize waiting less. Specifically, we are able to obtain information about different service types, from amusement parks to city sights. For any type, we also observe variations around the ticket price paid by consumers (free entry vs. not), as well as the service context, such as physical characteristics (indoor vs. outdoor), experience characteristics (active, visitor-involved interactions vs. passive, provider-directed interactions), human touch quality or location quality.

The challenge with multi-site evidence is to be able to identify congestion as a time-varying covariate. We obtain these attributes from text mining a large body of reviews, in the spirit of Mejía et al. (2021) or Hu et al. (2021). Specifically, congestion is measured at the day level, by looking at all the reviews around a certain date, and inferring the percentage of those that refer to queues and waiting. We also retrieve location features, e.g., price, from either direct sources, or from text mining through the appearance of certain themes as in Mejía et al. (2021). As an illustration, in Figure 1 we provide congestion dynamics (as measured by actual visitors and inferred from review comments) for the temple Angkor Wat in Siem Reap, Cambodia and the Victoria & Albert Museum in London, United Kingdom. We see in the figure that the reported incidence of congestion is closely related to the official number of visitors, suggesting that the inferred metric closely tracks actual congestion at the corresponding service site. Similar results were replicated in other sites in which monthly visit figures were available. Furthermore, we verify that the congestion metric is also highly correlated with the number of reviews posted, again suggesting that this seems to be a valid proxy for the congestion experienced on a given day. (Coefficients of correlation between visitor numbers and inferred congestion at the attractions in Figure 1 are 0.77 and 0.54, respectively; in comparison, correlations between visitor numbers and review numbers are 0.60 and 0.19; and correlations between review numbers and inferred congestion are 0.67 and 0.54.)

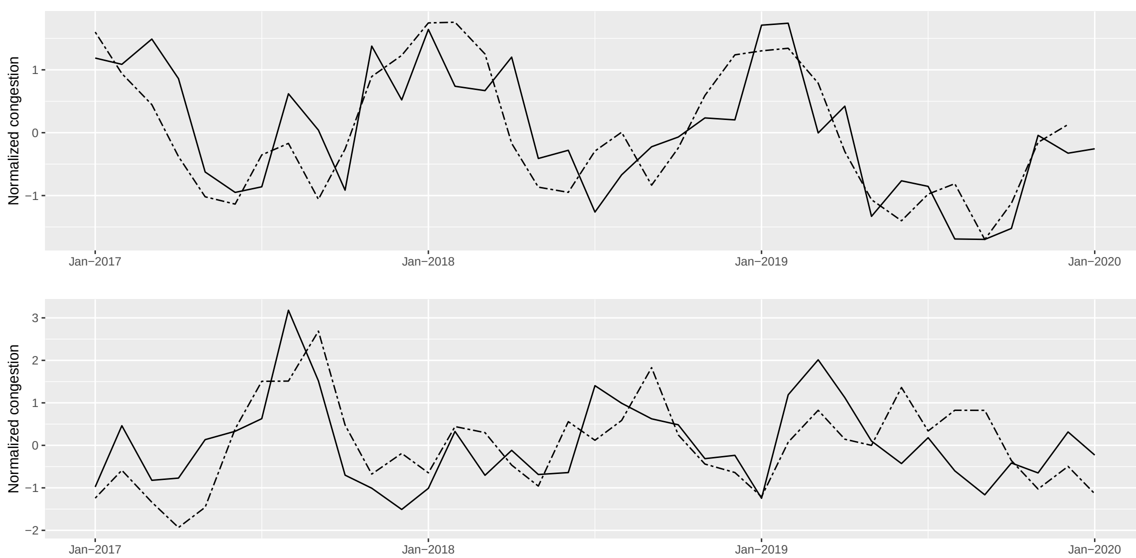

Figure 1 **Normalized congestion (z-score) at the Angkor Wat (top) and at the Victoria & Albert Museum (bottom) as measured by the o cial number of visitors (solid line) and reported average congestion (dashed line). Data is aggregated at the month level.**

Once attributes are obtained, we can study how reported consumer satisfaction is related to congestion and other service features. We build a reduced-form model in which the link between congestion and satisfaction depends on these service features. We estimate this model while controlling for user fixed effects, which heavily influence outcomes; this is possible because we have multiple ratings for the same user across various locations, in line with Martínez-de-Albéniz et al. (2020) and Deshmane et al. (2023). As expected, more congestion is generally associated with lower satisfaction. Beyond this intuitive finding, we are interested in unveiling the drivers of the sensitivity of satisfaction with respect to waiting. There is theory on this question (e.g., Larson 1987), but scarce empirical results validating or disproving the theory. The goal of this paper is precisely to explore such evidence.

We find that the nature of the service is the most important driver of consumer sensitivity to wait. Namely, sensitivity is highest in amusement parks, aquariums, and zoos and lowest in palaces, castles, places of worship, city areas, and ancient ruins. For instance, on average and other things equal, on a day on which 10% of visitors reported congestion, the probability that a given individual reports an excellent experience (i.e., rates the visit with 5 stars out of 5) at a zoo is 6.7% lower than when there is no congestion, while it is 3.3% lower at a palace sight. This means that zoos are more sensitive to congestion, compared to palaces. Interestingly, this relationship is not driven by the site's average congestion level or the price paid at the attraction. The same is true for the physical characteristics of the location, and outdoor and indoor services have similar sensitivities. These are negative results that contrast with theoretical discussions about the importance of the spatial layout (Baker and Cameron 1996).

On the other hand, our results also identify variables that strongly influence sensitivity to waiting. For instance, we find that the quality of the human interactions and the physical premises, specifically the perception of employee attentiveness and the perception of cleanliness, reduce sensitivity by a significant amount. For instance, on average and other things equal, on a day on which 10% of visitors reported congestion, the probability that a given individual would report an excellent experience would be 2.8% lower in a clean palace vis-à-vis 4.4% in a dirty one.

Moreover, we find that sensitivity is also dependent on the user's ease of access to the attraction. Specifically, local visitors are significantly more sensitive to congestion, compared to tourists. This is an intuitive result, because locals may be able to choose when they visit the attraction, leading to regret if they face high congestion, when they could have visited it during low congestion periods.

Taken together, our results suggest that congestion typically reduces satisfaction, but this strongly depends on the service design. By improving seemingly-minor aspects of the service process, such as attentive staff and facilities cleanliness, a service provider is able to mitigate the negative impact of congestion. Furthermore, under this better-controlled process, the provider could admit more visitors, hence increase revenues without sacrificing service quality as perceived by visitors, and therefore improve longer-term word-of-mouth recommendation and retention.

The rest of the paper is organized as follows. We discuss the related literature in §2. Section 3 describes the visitor utility model and identification strategy, while §4 describes context and data. Section 5 reports our empirical findings. We discuss a case study on museums in §6 to study the impact of quality shocks. Section 7 concludes. Analysis details and supporting tables are included in the Appendix.

2. Literature Review

Our work is related to the literature on service quality, and specifically the impact of congestion, as well as to studies that have used online reviews to measure consumer experience.

Since the first measurements of service quality (Parasuraman et al. 1985), waiting has been identified as a negative factor. Taylor (1994) provides the first empirical evidence that longer waits are associated with lower satisfaction levels and higher uncertainty and anger, in an airline context.

Many factors influence the perception of congestion (Maister 1985). The seminal paper of Larson (1987) identifies social justice, the queueing environment and feedback as essential drivers of consumer utility while waiting. Katz et al. (1991) provides an early comparison of different feedback mechanisms in a bank context. Baker and Cameron (1996) dissect the different elements of the service environment and theorize about their impact. They provide an extensive theoretical discussion on the drivers of sensitivity to waiting, and our results support some of their hypothesis (e.g., P6 on furnishings, related to cleanliness; and P12 on social facilitation, related to attentiveness) while suggesting some others that are still waiting for validation. Houston et al. (1998) further structure these influences and provide empirical evidence in a bank context. In the field, Sulek and Hensley (2004) show that fairness in the order of seating leads to better experiences in a restaurant. In the lab, Rafaeli et al. (2002) study how queue structure and position impact consumer psychological states and Zhou and Soman (2008) highlight that having a FIFO queue discipline is more important than reducing variations of waiting times across users in a queue. Voorhees et al. (2009) also show that fairness, affective commitment and environment quality moderate the effect of waiting on anger and regret, in four service contexts. Finally, Bitran et al. (2008) provide an integrative view of the literatures in psychology, marketing and operations.

The impact of congestion goes beyond consumer satisfaction. Perdikaki et al. (2012) and Lu et al. (2013) show that longer queues reduce purchase probability in retail. Allon et al. (2011) associate longer waits to lower market shares in fast-food restaurants. Ülkü et al. (2022) show that, after waiting, a consumer may reduce its service rate, thereby generating negative externalities on others. Waiting can also increase the perception of value of an experience (Koo and Fishbach 2010, Kremer and Debo 2016), or increase the subsequent consumption (Ülkü et al. 2020).

Furthermore, the literature in Operations Research has studied how to manage queues, so as to maximize social welfare. The celebrated $c\mu$ rule, first established by Smith (1955) and Cox and Smith (1961), and later extended by many, e.g., Van Mieghem (1995), states that priority should be given to users with higher waiting costs c and service rates μ, which constitute priority rules different from fair, FIFO discipline. More recently, Allon and Zhang (2017) propose to allocate more resources to more influential consumers, those with higher degree of centrality, so they generate a more positive word of mouth. Furthermore, behavioral aspects in queues have also been modelled, see Allon and Kremer (2018) for a recent review.

Another important connection of our work is the large literature that uses online reviews. Usually, ratings from reviews have been used as measurements of quality, despite possible problems of censoring (Dellarocas et al. 2010), manipulation/fraud (Dellarocas 2006, Hollenbeck et al. 2019) or bias (Chen and Lurie 2013, Chen et al. 2021). A fraction of this literature links quality to higher performance (Ba and Pavlou 2002, Chevalier and Mayzlin 2006). Some other works focus on studying the antecedents of ratings themselves. Huang et al. (2016) show that past experiences of a user influence restaurant ratings. Martínez-de-Albéniz et al. (2020) use sequences of restaurant reviews to study reference effects at the user level, and how these influence user choices among the available options and the resulting satisfaction. Deshmane et al. (2023) show that satisfaction is infused with negative intertemporal spill-overs due to habituation and reference effect adjustments, but also positive intertemporal spill-overs due to assimilation effects. They provide evidence from books, movies, restaurants and touristic attractions (the latter is similar to the data used in this paper).

More recently, reviews have been used for new purposes, which are closest to this paper. Mejía et al. (2021) uses text mining of Yelp reviews to break down service quality into its different components, and link it to restaurant survival. In particular, they find that long wait times increases the probability of closure. Hu et al. (2021) also mine Yelp reviews to measure the impact of virtual queues on consumer congestion perception. Our work similarly uses reviews to provide a measurement of congestion on a given attraction and day, but then use it to measure the impact of congestion level on satisfaction. This process allows us to expand the link between congestion and satisfaction across different service environments in the tourism industry. In contrast, Mejía et al. (2021) and Hu et al. (2021) focus on restaurants only.

3. A Model for Individual-Level Satisfaction

3.1. A Reduced-Form Approach

In this section, we present a model that links visitors' satisfaction and congestion of attractions and discuss its appropriateness to explain the impact of the latter on the former. For an individual i, visiting a touristic attraction j, at time t, we can define the utility derived from the visit as

$$U_{ijt} = w_{ijt} + \varepsilon_{ijt}$$

with

$$w_{ijt} = \alpha_i + \alpha^0_{jq(t)} + (\beta + \gamma Z_{ijt}) X_{jt} + \alpha^{00} Z_{ijt} \tag{1}$$

In this general formulation, α_i captures a fixed effect related to visitor i's intrinsic preferences. A time-varying effect of attraction j's quality, $\alpha^0_{jq(t)}$, is also included, where $q(t)$ denotes a time window containing t during which quality remained unchanged, e.g., year or quarter. The latter is important to control for possible changes in entry prices, which may change yearly, or attraction content, e.g., an exhibition in a museum, a new roller coaster in an amusement park, or construction works. For instance, average star rating of Big Ben in London decreased from roughly 4.6 to 3.4 in 2017 right after the tower was partially closed for maintenance works. Note that the occurrence of these events is usually low and their duration long, typically quarters or years.

Our main variable of interest is X_{jt}, which measures congestion at attraction j on day t inferred from visitors' opinions. The coefficient β thus includes the base sensitivity to congestion. In addition, we consider moderators Z_{ijt}, which incorporate user-, attraction- and time-dependent variables that may influence the impact of congestion on utility. They will include static moderators, such as attraction type, whether access is free or whether the user is local, and dynamic moderators, such as whether staff was nice to visitors on date t. Furthermore, note that Z_{ijt} may also have a direct impact on utility, which may be partially but not fully captured by $\alpha^{0}_{jq(t)}$, the fixed effects of attractions; for this reason, we add the control $\alpha^{00} Z_{ijt}$. Finally, we let ε_{ijt} be the random shock to utility.

As utility is unobservable, we consider instead whether the visitor rated the experience as excellent, i.e., gave the attraction the maximum possible rating (e.g., five out of five in Trip Advisor). Specifically, we let U_{ijt} be a binary outcome, in which

$$U_{ijt} = 1 \text{ if and only if } U_{ijt} \quad \varepsilon^{0}_{ijt}$$

where ε^{0}_{ijt} is the (random) utility of an outside option that is taken as a benchmark for comparison. We assume that both ε_{ijt} and ε^{0}_{ijt} are Gumbel distributed, so that U_{ijt} follow a logit specification in which

$$Pr[U_{ijt} = 1] = \frac{e^{w_{ijt}}}{1 + e^{w_{ijt}}}$$

and can be estimated with standard statistical packages. We also replicate our results with normal shock distributions, which lead us to use a probit, see §5.7.3.

3.2. Identification

To estimate β and γ, the coefficients that link utility to congestion, we take advantage of variations of the congestion level at attraction j over time.

As we describe later in §4.2, we obtain measures of congestion from reports of individual visitors, which could raise the question of reverse causality, i.e., the visitor reported congestion because the experience was not good. To eliminate this concern, we measure daily *average* congestion across other visitors' opinions at individual attractions, so that the impact of any one visitor's review on the resulting congestion metric is very small. For instance, the average number of reviews used to create the variable X_{jt} in our sample is 40.4. Under these conditions, the congestion metric is essentially the report of congestion by others, which is independent from the focal visitor's shock ε_{ijt} when visitors are independent from each other. As a result, $cor(X_{jt}, \varepsilon_{ijt}) \quad 0$ and the model can be correctly estimated. Note that in the robustness section, we consider an alternative metric X_{ijt} which uses a metric that *excludes* the focal visitor's review in the computation of congestion, so the metric becomes visitor-specific. We find that the results are unchanged.

Another possible concern for identification is that X_{jt} may be confounded by unobservables, such as the visitor type mix being correlated with the congestion level, e.g., more foreign tourists when congestion is high. These confounders will be controlled for in two ways: through attraction-time effects, which control for seasonality in base utility, and through individual visitor fixed effects. In other words, identification will be obtained at each attraction by looking at variation of utility *within the same period, the same visit condition and the same user type* across different congestion levels. This approach is conservative, and implies, among others, that visitors that always report excellent satisfaction will be dropped in the estimation. This is the case for single-visit users for instance, which constitute close to 40% of the observations for the top-100 attractions (see Table 1), but this is an acceptable price to pay to properly control for visitor heterogeneity.

4. Application Context and Data

In this section we present the framework and main data set used in our analysis, explain how we extract a measure of congestion and related variables from visitors' reports, and perform a preliminary analysis of the link utility-congestion, which paves the way to a formal regression analysis in §5.

4.1. Data Description

We use data about attractions retrieved from Trip Advisor, a major platform for travel information. Attractions, in accordance with Trip Advisor's classification, include museums, amusement parks, natural spaces, urban areas, zoos, ancient ruins, notable buildings such as castles, palaces, places of worship, towers, etc. We do not include in our study restaurants and hotels, in which content and prices change frequently together with congestion, while in attractions price changes are rare. We confine ourselves to the period elapsed from January 2011 to February 2020. Before 2011 there were relatively few reviews reported and after February 2020 tourism came to a halt due to the COVID-19 pandemic (until 2022 it remained severely disrupted).

Using web crawlers, we collected public data from all visitor reviews, including star rating (an integer between one and five), comments about the experience (unlimited text written in any language), date when the review was submitted, and unique identifiers for both attractions and users. During the customary cleansing process, we removed duplicates and reviews identified as translations of original postings, either manual or automatic. Also, to facilitate the analysis of text, we restricted it to the six most common languages used in reports, covering more than 90% of total comments. English was the most used language (52.1% of comments), followed by Spanish (11.8%), Portuguese (9.6%), Italian (7.4%), French (6.9%), and German (3.2%).

We thus obtained a corpus of 22,631,938 opinions by 8,966,624 users on 14,350 attractions worldwide. Figure 2 shows the distribution of attractions across countries and Table 1 shows relevant statistics regarding attractions and visitors.

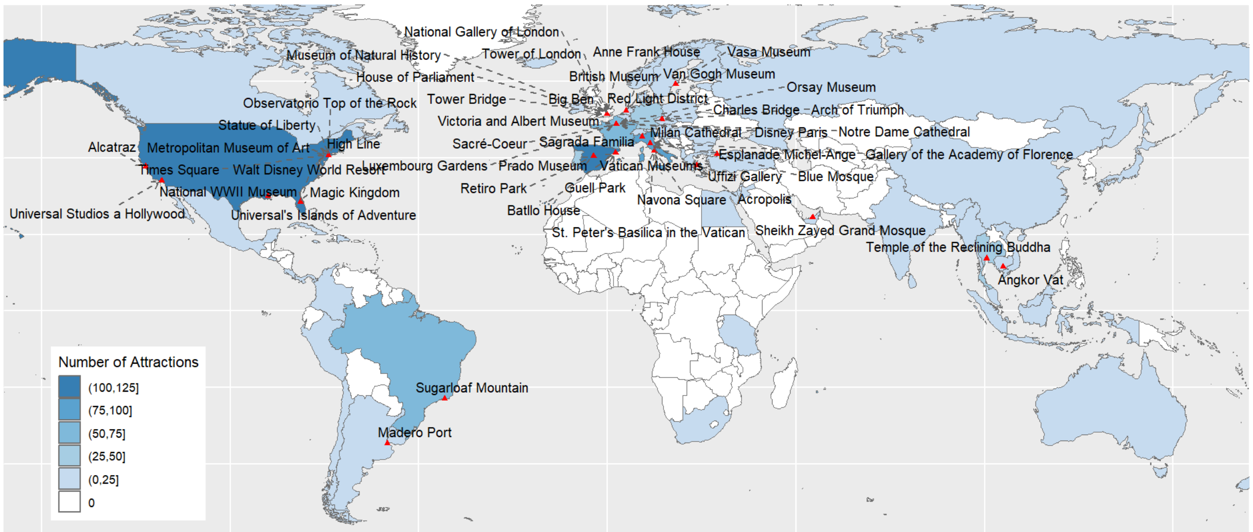

Figure 2 **Distribution of Top 500 attractions across countries. The names of top 50 attractions are included.**

Table 1 **Descriptive statistics.**

Attraction Popularity	Top 100	Top 250	Top 500	Top 1,000
attractions	100	250	500	1,000
unique visitors	1,775,346	2,746,247	3,588,184	4,300,038
visitors reporting 2+ sites †	32.4%	38.9%	42.8%	45.7%
reviews	3,123,767	5,868,507	8,768,806	12,004,597
reviews from visitors reporting 2+ sites	61.6%	71.2%	76.4%	80.4%
reviews per attraction (avg.)	31,238	23,474	17,537	12,005
reviews per attraction (coeff. of variation)	0.30	0.38	0.50	0.69
reviews per day (avg.)	9.33	7.01	5.24	3.59
reviews per visitor (avg.)	1.76	2.14	2.44	2.79
star rating (avg.)	4.56	4.53	4.51	4.49
excellent satisfaction (percentage rating 5/5)	68.1%	66.2%	64.8%	63.4%

†e.g., among visitors reporting on Top 250 attractions, 38.9% reported visits in 2+ attractions—out of 250.

The average visitor reports 2.5 visits, although there are large differences among visitors: 63.1% of them report a single visit, the third quartile is 2, and 10% report five visits or more. The distribution of star ratings in the resulting data is clearly skewed to the left: the mean score is 4.54. Roughly 68% of reviews rate the experience with a five out of five, the highest possible; the remaining percentages are 22.0 (fours), 6.5 (threes), 1.8 (twos), and 1.5 (ones). Such skewedness is one of the primary motivations behind our choice of excellent satisfaction as a binary outcome, following Martínez-de-Albéniz et al. (2020). Nevertheless, §5.4 replicates our analysis with alternative response variables such as good experience (whether the rating was four or five) and the rating itself.

As for attractions, Table 2 shows the distribution of the most popular ones per type. Types are assigned manually to the 1,000 attractions with the most reviews after unsuccessful attempts to do it automatically, due to inconsistent or incomplete taxonomies in popular sites, such as Trip Advisor or Wikipedia. The most common categories are city areas (e.g., Manhattan Skyline), museums (e.g., Anne Frank House), natural features (e.g., Copacabana Beach), places of worship (e.g., Blue Mosque), and buildings (e.g., Eiffel Tower). We removed from the tables including top lists the sites that can hardly be considered attractions, such as airport shuttles or tours agencies.

Table 2 Attractions per type (percentage in parenthesis).

	Top 100		Top 250		Top 500		Top 1,000	
amusement park	8	(8%)	20	(8%)	34	(6.8%)	52	(5.2%)
ancient ruins	2	(2%)	10	(4%)	17	(3.4%)	32	(3.2%)
aquarium	1	(1%)	5	(2%)	7	(1.4%)	14	(1.4%)
building	12	(12%)	23	(9.2%)	49	(9.8%)	84	(8.4%)
city area	27	(27%)	80	(32%)	156	(31.2%)	291	(29.1%)
museum	24	(24%)	45	(18%)	80	(16%)	162	(16.2%)
natural feature	4	(4%)	24	(9.6%)	55	(11.0%)	129	(12.9%)
palace/castle	5	(4%)	14	(5.6%)	31	(6.2%)	81	(8.1%)
place of worship	13	(13%)	21	(8.4%)	46	(9.2%)	99	(9.9%)
shopping area	1	(1%)	5	(2%)	18	(3.6%)	35	(3.5%)
zoo	3	(3%)	3	(1.2%)	7	(1.4%)	21	(2.1%)

4.2. Variable Construction

To evaluate individual congestion reported among visitors, we built an *ad hoc* filter, in which anchor key words—such as queue, wait, crowded, or packed—were identified in the text. Sentences containing these or similar words were further scrutinized using modifiers, i.e., additional keywords in the vicinity of anchors that modulate the meaning of the sentence. For instance, in the sentence "We were happy to find that the queue at the entrance was really fast", anchor key word "queue" and modifier "fast" are combined to conclude that there is no negative sentiment about congestion. We repeated this process for all six languages (Appendix B includes the details). As a result, if one review was identified as "negative"—i.e., expressing some sort of complaint about congestion—a binary variable C_{ijt}, was assigned value one, and zero otherwise.

To assess the performance of the filter, we randomly sampled 600 comments to find out that they were correctly classified in 80.2% of the cases. The remaining percentage was distributed between neutral comments (14.9%, as in, e.g., "The queue was not long.") and positive ones (4.9%, as in, e.g., "It was great that we didn't have to wait much."). The average attraction is reported to be congested in 6.5% of the cases, although percentages widely vary across attractions: for instance, average reported congestion in popular amusement parks is 13.5%, a number that grows to 24% in two sites: Disneyland Park in Anaheim, CA, and Alton Towers Resort in Staffordshire in the United Kingdom.Table 3 shows average congestion reported for the most popular attractions.

We use daily averages per attraction rather than individual visitors' congestion measures. A complication that arises is that comments are not necessarily submitted to Trip Advisor on the day of the visit. To get around this difficulty, we accounted for reported congestion not only on the current date—i.e., when the review was created—but also ρ days before and after, so that more individuals visiting the attraction on the same day are bound together. We use $\rho = 3$ as a plausible choice, implying that reported congestion is calculated using reviews in the week around the day of interest (see Appendix E for details). Formally, our congestion metric is constructed as

$$X_{jt} = \mathsf{E}[C_{i'jt'} jt \quad \rho \quad t^0 \quad t+\rho]$$

where the average is taken over all users i^0 that provided a review about j during that time window. Note that we discuss other values of ρ in §5.7.5.

Table 3 Percentage of visitors reporting congestion.

Attraction Popularity	Top 100	Top 250	Top 500	Top 1,000
amusement park	15.4	13.9	13.7	13.4
ancient ruins	10.0	7.5	6.7	7.1
aquarium	4.0	9.4	8.5	9.2
building	7.3	8.5	7.7	7.8
city area	7.9	7.5	7.0	6.5
museum	10.5	9.3	8.6	7.8
natural feature	5.2	7.4	8.2	7.6
palace/castle	13.6	9.9	8.8	7.4
place of worship	11.1	10.7	9.2	8.0
shopping area	6.2	5.4	6.0	6.0
zoo	5.4	5.4	6.3	6.5
All types	9.6	8.9	8.2	7.7

Similar filters are defined for staff attentiveness and facility cleanness by identifying key words related to, respectively, rudeness and dirtiness (see details in Appendixes C and D). Average reported rudeness and dirtiness for the Top 1,000 set are, respectively, 0.20 and 0.61 per cent, much lower than reported congestion. Note that, as in the case of congestion, we look for negative aspects of potential service drivers, e.g., we look for words related to staff rudeness, not staff attentiveness; dirtiness, not cleanness; and congestion, not the lack of it.

Prices for Top 1,000 attractions are current (as of July 2022) daily regular prices for adults measured in US dollars retrieved from the website if available, if not from Trip Advisor. Almost 44% of attractions are free, although this percentage varies across types: Less than 10% of amusement parks, zoos, aquariums, ancient ruins, and palaces/castles are free; while most city areas and all shopping areas are. The average ticket for the non-free attractions is \$24. Within these, amusement parks at \$72 on average are the most expensive, followed by aquariums (\$43), zoos (\$30), buildings, natural features, palaces/castles, and ancient ruins (\$22), museums (\$17), and city areas and places of worship (\$11).

We also obtain information about the visitor—specifically, the reported home location of the visitor, which allows us to identify those that are local (i.e., those whose home coincides with the attraction's location)—and the visit—whether the review is reported to happen with the visitor's family or not. We include binary variables that indicate if a visitor is a local (*is.local*) or if a visit is with family (*is.family*) in our analysis.

Table 4 includes descriptive statistics for the main variables used in the subsequent analysis, for the Top 500 attraction sample.

Table 4 Descriptive statistics for the Top 500.

	average	std dev	min	Q1	Q2	Q3	max	NAs
$congestion_{jt}$	0.082	0.095	0	0.03	0.07	0.12	1	0
$rudeness_{jt}$	0.002	0.022	0	0	0	0	1	0
$dirtiness_{jt}$	0.006	0.023	0	0	0	0	1	0
$price_j$	17.33	24.89	0	0	6.00	25.00	179	0
$is.local_{ijt}$	0.055	0.229	0	0	0	0	1	2,135,436
$is.family_{ijt}$	0.251	0.434	0	0	0	1	1	163,970

4.3. Model-Free Evidence

Before we present our model to analyze the relationship between congestion and satisfaction, we provide some evidence of the link between the two variables in Figure 3, in four well-known attractions: Sagrada Familia in Barcelona, the British Museum in London, Times Square in New York, and Disneyland Paris. The proportion of excellent experiences decreases in average congestion in all four attractions, but the slopes of the curves differ: for instance, if reported congestion grows from 0.1 to 0.2, the proportion of fives at Disneyland Paris decreases roughly by 0.05 points, while at Sagrada Familia it does so by a mere 0.01 points. The difference in slopes can be explained by differences of visitor mix and attraction features, as we will see in the next section.

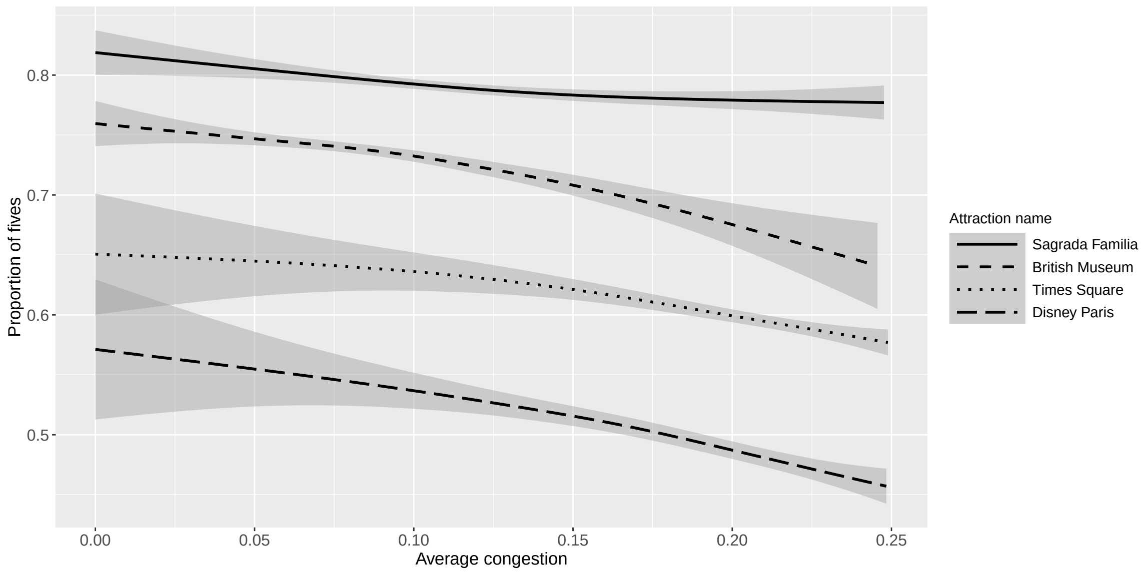

Figure 3 **Proportion of excellent experiences versus average daily reported congestion showing 95%-con dence bands, for selected attractions.**

5. Empirical Results

In this section, we derive our main empirical findings. We focus our study on the Top 500 attractions, which ensures enough reviews per day, but the analysis is replicated for the other three top lists, with similar results. We estimate Equation (1) for six different variants of interactions. Namely, we compare two models without congestion (only users' and attraction fixed effects in model 1 and users' and attraction seasonality effects in model 2), to models with congestion, with varying degrees of complexity: model 3, which assumes the same sensitivity to waiting across sites; model 4, with feature-driven sensitivity; model 5, with moderation effects of service quality; and finally, model 6, where each site has a different sensitivity.

In all cases, we estimate a logistic regression with a large number of fixed effects. All models contain fixed effects for users (α_i), those who reported two or more visits (5.59 on average) within the Top 500 list and whose ratings exhibit some spread. Note that we explore in §5.6 a more compact formulation with a one-dimensional continuous control of individual bias that delivers similar results. In addition, there are 500 fixed effects for attractions in the first model (α_j^0) and 17,915 attraction-quarter dyads in the remaining models ($\alpha_{jq(t)}^0$, with $q(t)$ denoting the year-quarter). As the number of fixed effects is very large, we resort to using the FEGLM package in R (short for fixed-effect generalized linear model), which is specifically designed to handle those, see Berge (2018) for details.

5.1. Congestion

Table 5 displays the estimates of the coefficients of models 1 to 5 together with goodness of fit measures. Note first that, even though the number of regressors increases with the complexity of

Table 5 Top 500 attractions. Coe cient estimates in models 1 to 5 (standard errors in brackets).

Model	1	2	3	4	5
Covariates					
$congestion_{jt}$			–0.82 (0.03)*** †		
$congestion_{jt}$ $amusementpark$				–1.37 (0.11)***	–1.24 (0.16)***
$congestion_{jt}$ $ancientruins$				–0.50 (0.20)*	–0.53 (0.24)*
$congestion_{jt}$ $aquarium$				–1.81 (0.32)***	–1.68 (0.33)***
$congestion_{jt}$ $building$				–0.94 (0.10)***	–0.87 (0.13)***
$congestion_{jt}$ $cityarea$				–0.51 (0.06)***	–0.57 (0.12)***
$congestion_{jt}$ $museum$				–0.94 (0.08)***	–0.83 (0.11)***
$congestion_{jt}$ $naturalfeature$				–0.91 (0.11)***	–0.95 (0.17)***
$congestion_{jt}$ $palace=castle$				–0.91 (0.11)***	–0.80 (0.15)***
$congestion_{jt}$ $placeofworship$				–0.52 (0.10)***	–0.46 (0.12)***
$congestion_{jt}$ $shoppingarea$				–0.74 (0.18)***	–0.70 (0.18)***
$congestion_{jt}$ zoo				–2.46 (0.39)***	–2.43 (0.41)***
$\mathbf{1}frudeness_{jt} > 0g$ z					–0.02 (0.01)*
$\mathbf{1}fdirtiness_{jt} > 0g$					–0.04 (0.01)***
$congestion_{jt}$ $\mathbf{1}frudeness_{jt} > 0g$					–0.25 (0.07)***
$congestion_{jt}$ $\mathbf{1}fdirtiness_{jt} > 0g$					–0.28 (0.06)***
$congestion_{jt}$ $notfree_j$					–0.01 (0.09)
$congestion_{jt}$ $outdoors_j$					0.15 (0.11)
Fixed effects					
User	Yes	Yes	Yes	Yes	Yes
Attraction	Yes	No	No	No	No
Attraction-quarter	No	Yes	Yes	Yes	Yes
number of observations	4,741,612	4,741,133	4,741,133	4,741,133	4,741,133
number of regressors	848,301	865,715	865,716	865,726	865,732
pseudo-R2 (McFadden)	0.54427	0.54961	0.54968	0.54969	0.54973
AIC	6,880,373	6,854,513	6,853,732	6,853,626	6,853,134
^{2}-test p-value (current vs. previous, e.g., 2 vs. 1)		$< 10^{-16}$	$< 10^{-16}$	$< 10^{-16}$	$< 10^{-16}$

†Significance codes: 0 ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’.’ 0.1 ’ ’ 1 ‡$\mathbf{1}f\ g$ represents the indicator function.

the model function, the quality of the model improves as measured by its AIC function.

The coefficient of average congestion in model 3 is 0.82 (with standard error 0.03), which shows that rating decreases—or more exactly, departs from five—with reported congestion. Specifically, if the latter increased by one percentage point, from 8.2% to 9.2%, then expected percentage of fives would roughly drop by 0.2 percentage points (i.e., log-odds decreases from an average of 0.61 to 0.60).

5.2. Drivers of Sensitivity to Congestion from Attraction Characteristics

As for attractions, the coefficients of the products of congestion times the attraction types’ dummies for all eleven attraction types identified in model 4 are negative and significant, revealing that higher congestion leads to lower ratings for all types, something not obvious *a priori*, especially in cases such as city areas or natural features in which visitors do not have to line in a queue for access. Still, as the results show, they still prefer fewer people around to more, e.g., a deserted beach and a quiet city main square.

Also noteworthy is the fact that such coefficients differ from each other depending on the type of attraction, a result in line with evidence from Figure 3. Figure 4 illustrates these coefficients by showing 95%-confidence intervals around model 4's coefficients, ranked by type. Typical visitors of zoos, aquariums, and amusement parks see their satisfaction most reduced when facing congestion, controlling for users' biases, attractions' inherent quality and dynamics. In contrast, visitors at places of worship, ancient ruins, and city areas stand out as the most lenient: for the same congestion level, they penalize rating the least. Consider a palace and a zoo that share $\hat{\alpha}^{0}_{jq(t)}$ in a particular quarter: On a no-congestion day at both attractions, the probability that a given user gives a five is the same at the two attractions and the odds ratio palace-to-zoo is one. On a 10%-congestion day, the odds ratio will be $e^{0.1\,(2.46\;\;0.91)} = 1.168$. This means that the odds of scoring a five at the palace will be 16.8% higher than at the zoo. Given that the baseline probability of having an excellent experience is 0.66, this would imply that a 10% congestion level reduces this probability to 0.61 in the zoo and 0.64 in the palace.

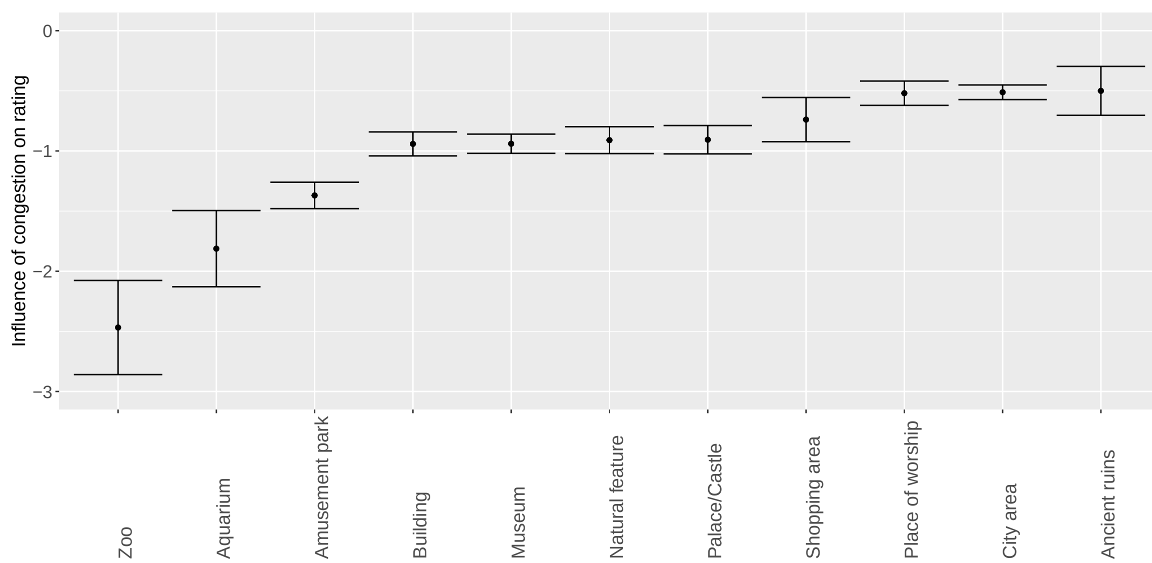

Figure 4 **Top 500 Attractions. Model 4's 95%-con dence intervals for coe cients, ranked by type. Dots denote averages.**

Differences across types may be explained in light of the findings by Taylor (1994), which in a service realm asserts that customers' anger increases when delays feel uncertain, when they are controllable by the service provider and their causes frequent. When considering these factors. the attraction types that are most sensitive to congestion, namely zoos, aquariums and amusement parks, are usually manned, relatively expensive, and visitors expect that the service provider exerts an effort to keep things under control, including ordinary queues. Furthermore, indoors, subject-to-uncertainty enclosed queues to enter the building or getting access to popular species may be annoying in the case of aquariums; open-air queues not sheltered from heat or rain, can reduce comfort in the case of zoos. In contrast, the attraction types that are least sensitive to congestion are usually unattended, free or unexpensive, hence there is a much lower expectation from visitors for the service provider to take care of congestion. Indeed, congestion hurts visitors less when the responsibility of the delay cannot be attributed to the service provider (hypothesis P9 in Baker and Cameron 1996).

5.3. Drivers of Sensitivity to Congestion from Visit Characteristics

Model 5 includes six additional covariates to model 4 to control for other potentially relevant features, which measure staff attentiveness (rudeness), facility cleanness (dirtiness), price (free or not), and location (outdoors or not). Rudeness and dirtiness are measured similarly to congestion, i.e., taking averages in the interval $[t - \rho, t + \rho]$. Since, as said, the incidence of these two variables is very low, we binarized these covariates using an indicator function, which yields one if the corresponding variable (e.g., $rudeness_{jt}$) is larger than zero, and zero otherwise. (We also run models taking logarithms of those covariates—e.g., $\log(rudeness_{jt} + \epsilon)$ for a small positive ϵ—with the same purpose and got similar qualitative results.) As Table 5 shows, type estimates only slightly change with respect to those of model 4 due to the inclusion of new covariates, which reassures our conclusions from model 4. Two observations stand out from model 5.

First, *ceteris paribus*, there is no evidence that price is relevant to explain sensitivity to congestion, as per the coefficient of *notfree* being non-significant (we also ran models with only *price* and both *price* and *notfree* and coefficients are still non-significant). One could anticipate that priced attractions, rather than free ones, would make visitors less lenient regarding congestion. Our results, however, suggest that, once the type of attraction is considered, this is not the case.

Second, other things equal and in contrast to price, the lack of staff attentiveness and location cleanness, two key operational variables in service settings, have a strong, negative impact on the sensitivity to congestion, as per the corresponding coefficients being significant. Even if attentiveness and cleanness are not related to congestion, visitors become more tolerant to queues if are treated politely and feel a spotless atmosphere around them. This is a statistical relationship that could be further studied in behavioral experimental studies to more clearly delineate the underlying psychological mechanism. The direct implications for the service managers are not only clear, but actionable: conscious efforts must be exerted by the staff to be more polite with visitors and to keep the site clean and organized. These actions have a double impact on rating: one direct—the obvious one—and one indirect, through higher lenience to congestion.

5.4. Drivers of Sensitivity to Congestion from Visitor Characteristics

Individual sensitivities play indeed a role in visitors' perception of congestion. As said, those are accounted for in users' fixed effects in model 5, and an alternative approach is presented in §5.6. Besides those, we address next the potential influence on sensitivity to congestion of two visitors' attributes, sometimes reported by users: whether they are local (i.e., come from the place where the attraction is located) and whether the visit is a family one. To do so, we modify model 5 to include four additional indicator variables in model 5': Table 6 show the results (output from model 5 is also shown for comparison purposes). Note that the number of observations and regressors is

Table 6 Top 500 attractions. Coefficient estimates in models 5 and 5' (standard errors in brackets).

Model	5	5'
Covariates		
$congestion_{jt}$ $amusementpark$	–1.24 (0.16)*** †	–1.18 (0.19)***
$congestion_{jt}$ $ancientruins$	–0.53 (0.24)*	–0.32 (0.28)
$congestion_{jt}$ $aquarium$	–1.68 (0.33)***	–1.56 (0.37)***
$congestion_{jt}$ $building$	–0.87 (0.13)***	–0.88 (0.15)***
$congestion_{jt}$ $cityarea$	–0.57 (0.12)***	–0.49 (0.14)***
$congestion_{jt}$ $museum$	–0.83 (0.11)***	–0.82 (0.13)***
$congestion_{jt}$ $naturalfeature$	–0.95 (0.17)***	–0.88 (0.19)***
$congestion_{jt}$ $palace=castle$	–0.80 (0.15)***	–0.70 (0.17)***
$congestion_{jt}$ $placeofworship$	–0.46 (0.12)***	–0.43 (0.14)**
$congestion_{jt}$ $shoppingarea$	–0.70 (0.18)***	–0.59 (0.21)**
$congestion_{jt}$ zoo	–2.43 (0.41)***	–2.53 (0.47)***
$\mathbf{1}\{rudeness_{jt} > 0\}$ ‡	–0.02 (0.01)*	–0.02 (0.01)*
$\mathbf{1}\{dirtiness_{jt} > 0\}$	–0.04 (0.01)***	–0.04 (0.01)***
$congestion_{jt}$ $\mathbf{1}\{rudeness_{jt} > 0\}$	–0.25 (0.07)***	–0.24 (0.07)**
$congestion_{jt}$ $\mathbf{1}\{dirtiness_{jt} > 0\}$	–0.28 (0.06)***	–0.24 (0.06)***
$congestion_{jt}$ $notfree_j$	–0.01 (0.09)	0.03 (0.11)
$congestion_{jt}$ $outdoors_j$	0.15 (0.11)	0.17 (0.13)
$is.local$		0.48 (0.01)***
$is.family$		–0.01 (0.01)
$congestion_{jt}$ $is.local$		–1.72 (0.15)***
$congestion_{jt}$ $is.family$		–0.18 (0.05)***
Fixed effects		
User	Yes	Yes
Attraction	No	No
Attraction-quarter	Yes	Yes
number of observations	4,741,133	3,721,642
number of regressors	865,732	865,736
pseudo-R2 (McFadden)	0.54973	TBD
AIC	6,853,134	5,314,772

†Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

‡$\mathbf{1}\{\}$ represents the indicator function.

lower than in model 5, as not all visitors disclose their origin and/or type of visit.

Families appear to dislike waiting more, as per the coefficient of $congestion_{jt}$ $is.family$ being negative and significant. This is expected indeed, especially when there are small children involved. Also, locals are much harsher when it comes to congestion, as per the coefficient of $congestion_{jt}$ $is.local$ being negative, very large in absolute value, and significant. We discuss this notable finding further in §6.1. These results support hypotheses H1c in Houston et al. (1998) and H6 in Antonides et al. (2000), which claim that waiting cost has a negative impact on the perception of waiting.

Furthermore, our results suggest that is not only the direct waiting cost that impact perception—as the cost of a telephone call while waiting, as in Antonides et al. (2000), but also the opportunity cost due to waiting. This is the case of locals that visit a museum, who may have tighter schedules than tourists or may experience regret knowing that they could have visited the attraction on a different, less-congested date. Our finding suggests that local visitors should be discouraged from visiting attractions on congested days, by offering discounts on non-busy days or informing them about which days are busier. Not only persuaded locals will enjoy the attraction more, but tourists will benefit as well, as congestion of busy days will decrease.

To summarize the results so far, we can offer four main insights about the antecedents of the sensitivity to congestion. First, sensitivity is extremely dependent on the nature of the attraction, suggesting that service managers have limited degrees of freedom to influence visitor tolerance to congestion once the attraction type has been determined. Second, price or the physical premises have little influence on the sensitivity, suggesting that it is not directly clear how one could use monetary incentives to directly compensate consumers for the disutility from congestion. Third, the service provider can use other service delivery operational variables—staff attentiveness and cleanliness—to make visitors more tolerant to congestion. Fourth, locals are much more impatient with queues than their non-locals counterparts, and should then be managed differently.

5.5. Taxonomy for Individual Attractions

We next make use of model 6 to provide additional insights for individual attractions, as it includes a separate sensitivity to congestion β_j for each attraction.

Figure 5 presents a taxonomy for the Top 500 attractions, based on four dimensions: the attraction type, the proportion of visitors reporting congestion (X axis), the coefficient estimate $\hat{\beta}_j$ of the attraction in model 6 (Y axis), and its significance (solid dots are significant at the 95% level). For instance, the solid point that appears to the southwest in the "building" panel corresponds to the Tower Bridge in London: about 4% of the visitors to this attraction reported congestion and the coefficient estimate of variable $congestion_{jt} \quad tower\_bridge$ in model 6 is equal to -3.2 and significant (p-value < 0.05).

The graph unveils a number of relevant facts. First, with a few exceptions (eight cases out of 122, or 6.6%), all significant coefficients are negative. This fact reinforces the general finding that visitors penalize congestion. Second, with a few exceptions (noteworthy is the case of museums, which may indicate the existence of non-linear effects), sensitivity to congestion does not substantially change with average reported congestion. (Table 7 shows, for each attraction type, the coefficient δ in regression $\beta_j = \alpha + \delta \quad avg.congestion_j$). This is especially the case for amusement parks, palaces, castles, and natural features, where solid dots are located on horizontal bands. And third,

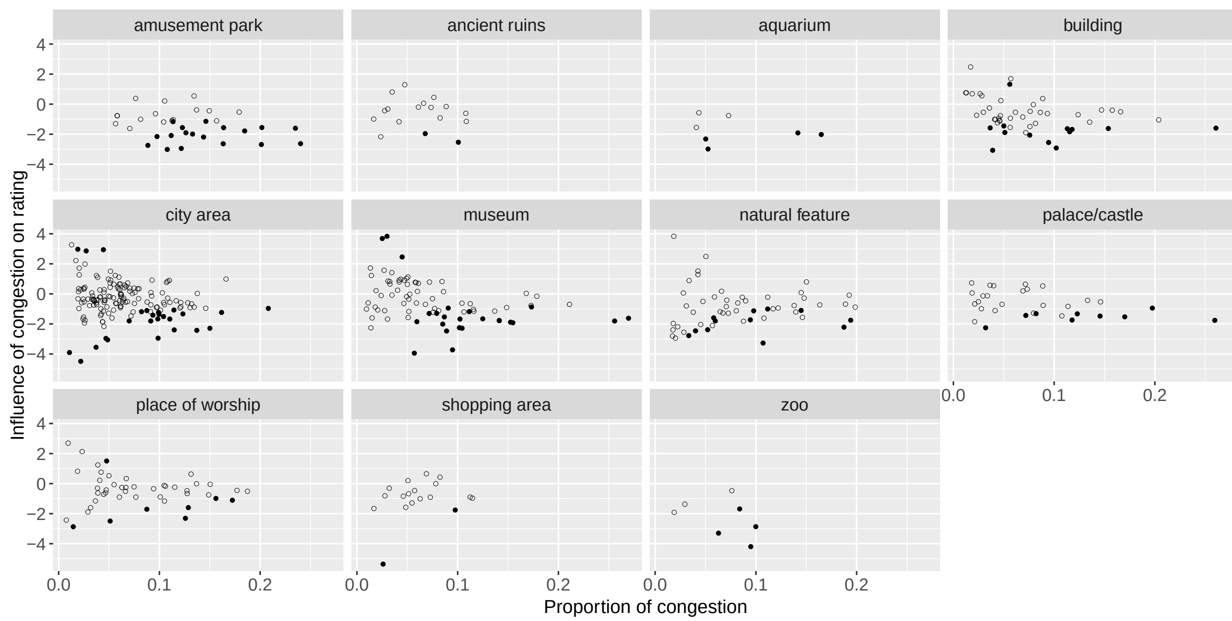

Figure 5 **Top 500 Attractions. Model 6's coe cients for congestion of individual attractions (the** $\hat{\ }_j$ **of the product** $congestion_{jt}$ $attraction_j$**) vs. average proportion of reported congestion, per type. Solid dots represent signi cant coe cients at 95% con dence level.**

Table 7 **For each attraction type, sensitivity of individual betas to average congestion incidence** $avg{:}congestion_j$ **(standard errors in brackets).**

amusement park	−5.29 (3.33)
ancient ruins	−5.06 (8.69)
aquarium	−2.56 (7.38)
building	−6.94 (2.99)*†
city area	−6.23 (2.60)*
museum	−10.03 (2.65)***
natural feature	0.74 (3.44)
palace/castle	−5.35 (2.47)*
place of worship	−3.19 (3.42)
shopping area	13.66 (10.9)
zoo	−15.72 (16.5)

†Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

there are large differences among attractions within types, signaling that there are other variables (recall, outside visitors, inherent attraction quality, and slow attraction dynamics already captured by fixed effects) which may impact visitors' perception of congestion. This fact allows for the management of individual attractions to benchmark their relative position in comparison with their peers. For instance, the management of the Tower Bridge may ask why visitors are so severe regarding congestion compared to other buildings, even when reported congestion is among the lowest in that group. In contrast, management of the Lello Bookstore in Porto, Portugal, may feel pleased that, even when 21% of visitors report congestion, this does not harm its rating much.

Similarly, as Figure 6 shows, coefficients are insensitive to the quality of the attraction, as measured by the average of the fixed effects of attractions. (Table 8 shows, for each attraction type,

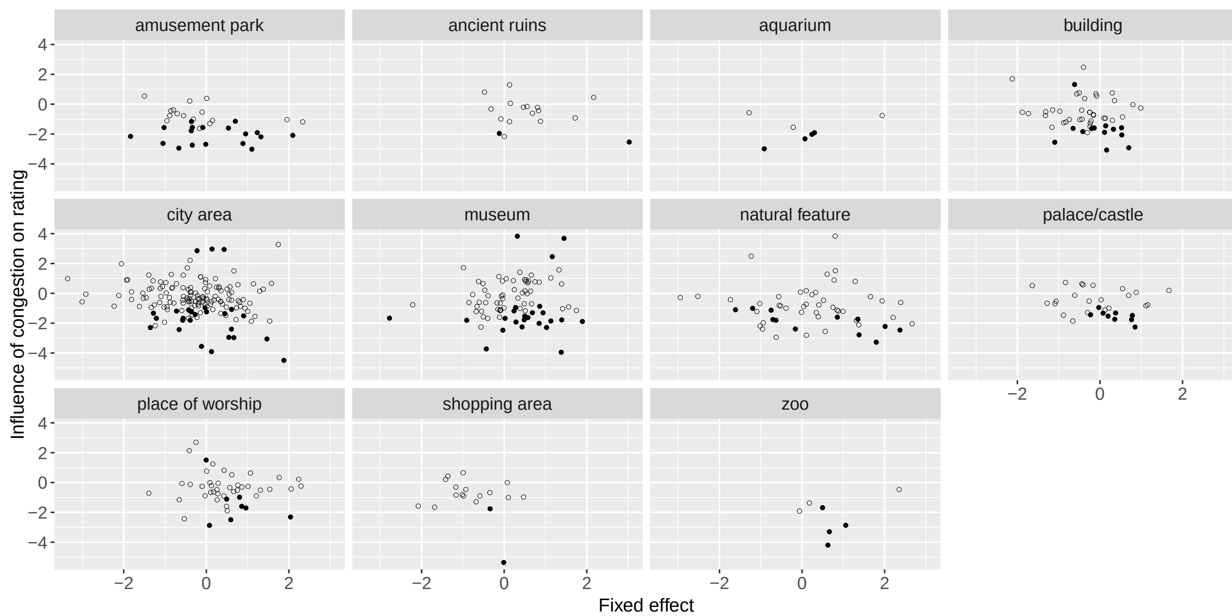

Figure 6 **Top 500 Attractions. Model 6's coefficients for congestion of individual attractions (the $\hat{\beta}_j$ of the product $congestion_{jt} \times attraction_j$) vs. attraction fixed effects, centered around the mean, per type. Solid dots represent significant coefficients at 95% confidence level.**

the coefficient δ in regression $\beta_j = \alpha + \delta \cdot fixed.effect_j$). In sum, once type is accounted for, the

Table 8 **For each attraction type, sensitivity of individual betas to $fixed{:}effect_j$ (standard errors in brackets).**

amusement park	–0.23 (0.16)
ancient ruins	–0.29 (0.27)
aquarium	0.21 (0.36)
building	–0.35 (0.23)
city area	–0.20 (0.10).†
museum	0.19 (0.20)
natural feature	–0.14 (0.15)
palace/castle	–0.22 (0.19)
place of worship	–0.15 (0.21)
shopping area	–0.48 (0.46)
zoo	0.56 (0.67)

†Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

sensitivity of rating with respect to congestion is by and large neither a function of the average congestion or the average rating of the attraction.

5.6. Accounting for Visitor Heterogeneity

In our base models, we account for visitor heterogeneity through fixed effects. This introduces complexity in the estimation process and forces us to remove many observations from visitors without variation in their reported ratings, e.g., those with a single report. We explore here an alternative approach to control for visitor biases.

5.6.1. Extent of Visitor Biases. We first assess the impact of removing visitor fixed effects. For this purpose, we run a modified version of model 6 as if users' identifiers were not available. That entails not including model 6's roughly 850,000 users' fixed effects and including about four million observations that had been removed in model 6 because of lack of variation in users' ratings.

The coefficient of correlation of the two arrays of 500 coefficients of the products $congestion_{jt} \times attraction_j$ of each model is 0.78; half of the corresponding coefficients differ by 53% or more; and 75 coefficients change from positive to negative or vice versa. Furthermore, the average of 500 coefficients in model 6 was -0.68, compared with -0.78 in the modified model. This suggests that users who wrote only one review within the Top 500 attractions are, on average, less tolerant to congestion than otherwise.

This finding shows that there is a substantial difference between the two models when assessing the rating sensitivity to congestion. It underscores the importance of considering visitor individual biases in the authorship of reviews, which is possible in online feedback systems, but not in other traditional quality measurement methods, such as on-site anonymous surveys.

5.6.2. Alternative Controls for User Biases. A difficulty we encountered when using fixed effects for users is that the number of variables in the models grows tremendously, making models non-parsimonious and, at times, computationally unfeasible. For instance, we were not able to run model 6 with 1,000 attractions in our server due to lack of memory to handle large arrays. As an alternative, we modify model 6 to account for users' biases in a more compact way: we define a new variable, $optimist_i \in [0,1]$, defined as the proportion of reports with excellent satisfaction (rating five out of five), given by the same user across all visits, possibly including those in other contexts not related to the current study, such as restaurants or hotels. This variable indicates how positive the user generally is in comparison with other users. We then define the log-odds of this variable as $bias_i := \log\left(optimist_i/(1 - optimist_i)\right)$ and include a control $\hat{\beta} \times bias_i$ in Equation (1) replacing the fixed effect α_i. Note, when using $bias_i$, users without variation in ratings, i.e., those with $optimist_i \in \{0,1\}$, are dropped, thus leading to the data sets of original and modified models having the same observations.

The estimates of all 500 products $congestion_{jt} \times attraction_j$ are similar to those of model 6 as measured by the coefficient of correlation, which is 0.92. Furthermore, removing almost 850,000 variables (the visitors' fixed effects) from the model improves its goodness of fit —e.g., AIC is reduced from 6,853,551 to 6,440,665. Therefore, using a single covariate rather than one fixed effect per user emerges as a plausible alternative to account for users' biases. Resorting to this proposed alternative arises as a viable solution when analyzing a single attraction, in which visitors report only one visit, but provide reports from other attractions not being studied. We make use of this approach in §6.

5.7. Robustness

In this section, we provide robustness checks, including alternative definitions of the notion of excellent experience, a probit specification, a linear functional link between congestion and satisfaction, and alternative measurements of the congestion variable.

5.7.1. Simple Specifications. Just for robustness purposes, we first include three simple models, namely 0, 0', and 1', consisting of various combinations of fixed effects for users and attractions, as well as, possibly, congestion. Table 12 in Appendix A shows that all metrics considered are consistent with previous results (output from models 1, 2, and 3 are also included for comparison purposes).

5.7.2. Definition of Excellent Satisfaction. So far, we have considered that a visitor had an excellent satisfaction when the star rating was five out of five. Here, we examine an alternative binary response variable, in which satisfaction is considered to be excellent when rating is four or five. By doing so, the proportion of excellent satisfaction grows from 0.68 to 0.90.

As Table 13 in Appendix A shows, the regression of model 3 that uses the modified response variable returns a negative, significant coefficient for congestion at 1.02, close to, but lower than the coefficient 0.82 in the original model 3 (see Table 5). This indicates that visitors with an intermediate satisfaction report are more sensitive to congestion. In other words, many visitors report the maximum star rating of five no matter the level of congestion, and a small fraction would move their rating to four when they experience congestion. In contrast, visitors who in principle would express a rating of four in the absence of congestion are more likely to report a three or less on congested days.

We also repeat the process with models 4 and 5. Figure 7 shows the coefficients of the original model 4 (those from Figure 4) and modified versions. The latter coefficients are quite close to the original ones, although most of them are larger in absolute value, in line with the above discussion about model 3. The gap between corresponding points in Figure 7 is larger for aquariums and zoos, revealing that the mentioned effect is larger for these attractions: not only congestion is penalized the most, but rating drops substantially more than in other attractions when facing congestion.

5.7.3. Probit Model. Our base model is based on Gumbel-distributed shocks, leading to a logit specification. We replicate our analysis with normal shocks and the probit specification. The results are included in Table 14 in Appendix A, which should be compared to Table 5. As we can see, the results are very similar—except for the expected fact that probit coefficients are smaller in absolute value, as the probit function has lighter tails. This shows that our results are robust to distributional assumptions on the residuals.

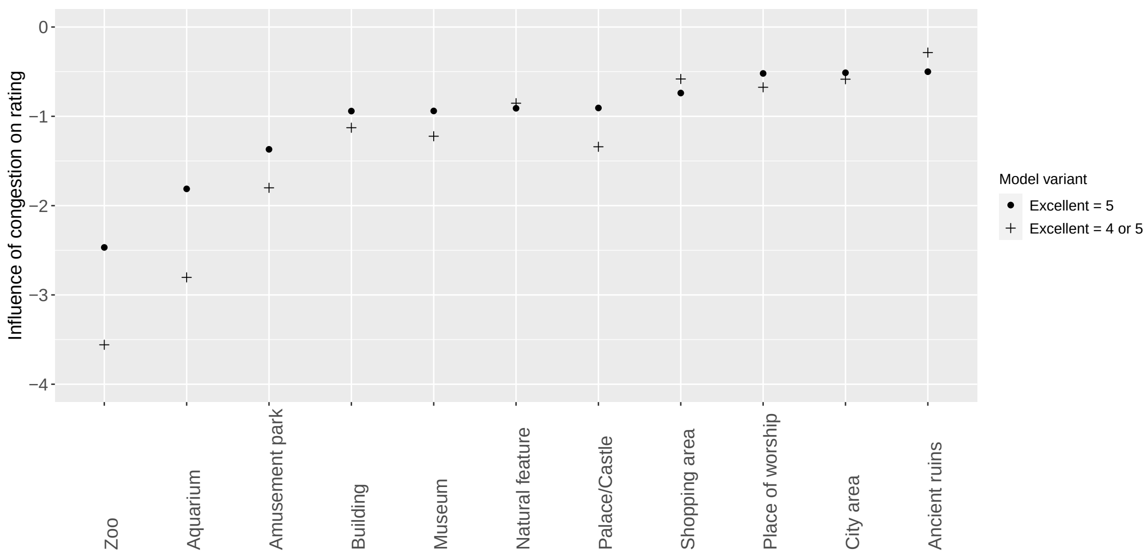

Figure 7 **Top 500 Attractions. Coe cients of model 4, original (excellent = 5) and modi ed (excellent = 4 or 5).**

5.7.4. Linear Model. Our choice of a two-level discrete response variable was driven by the large proportion of five star ratings, which induced us to focus on the difference between an excellent experience report and the rest. We now modify the response variable into a continuous response directly equal to the numerical star rating. We thus run specifications 3 to 5 using a linear regression model, instead of a logistic regression (see Table 15 in Appendix A).

The coefficient of modified model 3 is 0.22, again significant. The estimate can be shown to be consistent with the one from the original model. For instance, when the proportion of reported congestion increases from 0 to 0.1, rating in the modified model would drop by 0.022 on average, consistent with a reduction in probability of fives of 0.03 in the original model.

The coefficients of modified model 4 are shown in Figure 8 together with those of the original model. As we can see now, not only all coefficients are negative, but the relative position of coefficients within models is approximately the same; relatively large (small) coefficients in the original model are associated with relatively large (small) coefficients in the modified model.

5.7.5. Reporting Lag. As explained, we chose $\rho = 3$ (plus/minus three days from the day of interest) to compute average congestion. We explore here alternative values for the reporting lag. Figure 9 shows the normalized coefficient of model 3 for various values of ρ. Note that we normalize congestion—using z-scores—so that the variance of the resulting covariate, normalized congestion, coincide for different values of ρ , which allows us to make a fair comparison between the coefficients. As shown in the figure, changing ρ does not have a large impact on the coefficient:

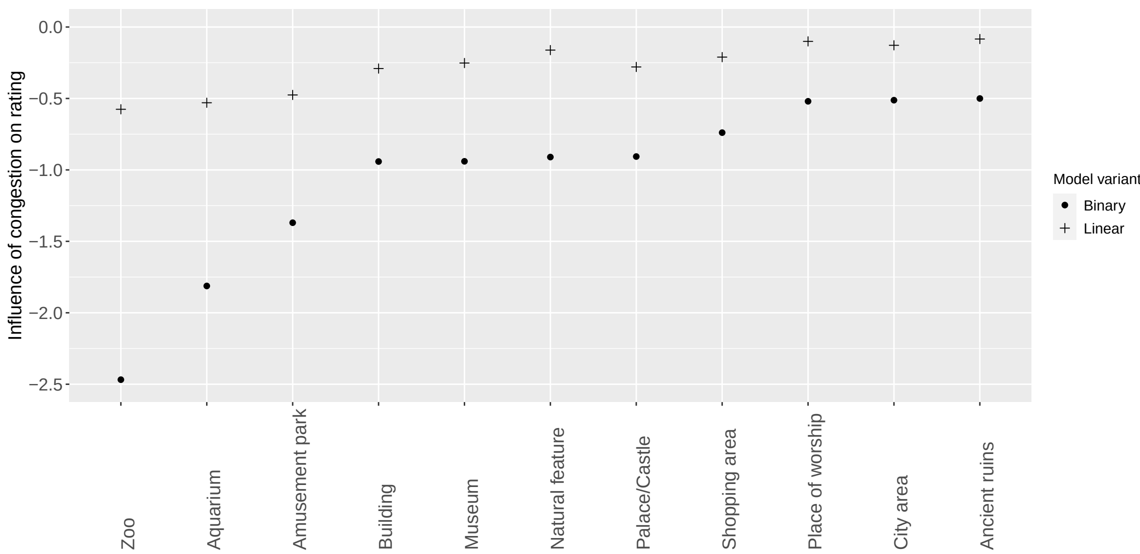

Figure 8 **Top 500 Attractions. Coe cients of model 4, original (logistic regression) and modi ed linear regression.**

points appear on a horizontal, narrow band. This suggests that our findings are robust to alternative constructions of the congestion variable.

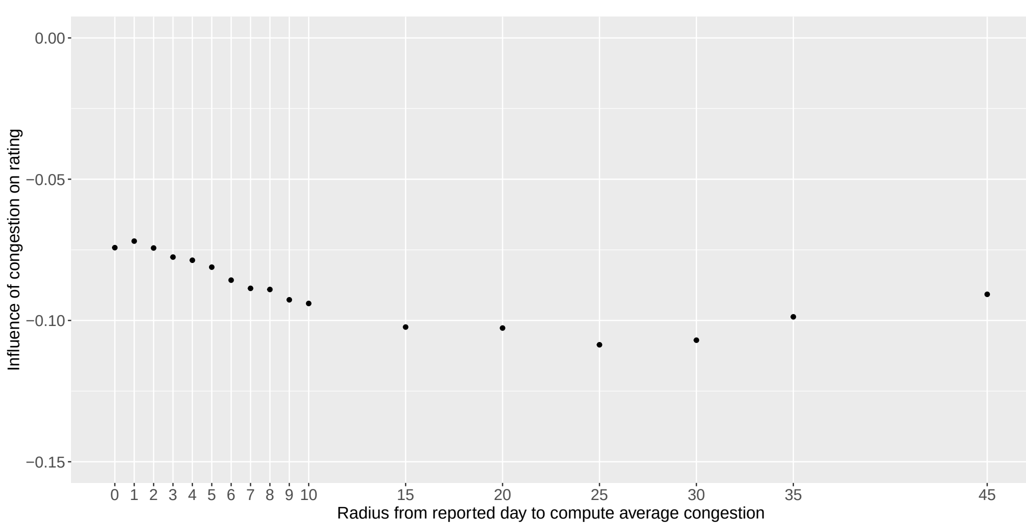

Figure 9 **Top 500 Attractions. Coe cients of model 3 for various values of . Solid dots denote signi cance at 95% con dence level.**

5.7.6. Alternative Measure of Congestion. As noted in the introduction, we measure daily average congestion across visitors' opinions, so that the impact of a particular visitor's review on the congestion metric is very small. We go now one step further by using an alternative metric

X_{ijt} that excludes the focal visitor's review in the computation of average congestion, so the metric becomes visitor-specific. By doing so, we completely rule out the chance that our results are driven by reverse causality. Table 16 in Appendix A shows the coefficient estimates for modified models 3, 4, and 5 when the alternative metric is used.

The results are qualitatively identical. The signs of all significant coefficients are preserved and the ranking of coefficients across attraction types does not essentially change. The absolute values of the correspondent coefficients of congestion are smaller than in the original models, something expected given that the focal individuals' opinions—eliminated from the calculations in the modified models—substantially contributed to reduce the coefficients of congestion in the original models.

6. Drivers of Sensitivity to Congestion at Museums

After the general analysis, we focus on one particular attraction type, museums. By narrowing our scope, we seek two goals. First, we explore whether the impact of key drivers of sensitivity to congestion in museums (attentiveness, cleanliness, local visitors) remains similar within this type of attraction compared to the rest. This will help us comment on the robustness of the results. Second, we test if a positive quality shock—a potential discretionary decision made by museum managers—make visitors more tolerant to congestion.

6.1. Staff Attentiveness, Cleanness, and Local Visitors' Perception.

We first complement model 5 (with user fixed effects) with a model 5" which includes visitors' biases using the method described in §5.6. We consider only the observations in which the attraction type is a museum. By considering the alternative method, we are able to obtain results at individual attractions too.

As for visitor outcomes, Table 9 presents the results for museums in general and then at selected museums. As the coefficients of $congestion_{jt}$ show, congestion remains as an important driver of dissatisfaction. More interestingly, we confirm—as in model 5—that the lack of cleanness in the average museum increases the intolerance of visitors to congestion. At the individual level, we also found several examples of dirtiness worsening congestion perception, which reinforces our general recommendation for managers to keep high standards of cleanliness to reduce the negative perception of congestion.

At the same time, we found staff attentiveness to be irrelevant in museums, which is intuitive because museums tend to be self-serviced environments, in contrast to attractions in which employees have a more important role, such as amusement parks.

As for visitor characteristics, to analyze the perception of locals compared to non-locals, we select British and Prado museums, which have the highest number of opinions (more than 2,000 each)

Table 9 Museums in Top 500. FEGLM Coefficient Estimates (standard errors in brackets).

	Museums	British M.	Louvre	MoMA	Prado	Van Gogh
model specification	5	5"	5"	5"	5"	5"
$congestion_{jt}$	–0.67 (0.28)** †	–0.96 (0.41)*	–1.43 (0.42)***	–0.26 (0.43)	–1.05 (0.45)*	–1.33 (0.31)***
$\mathbb{1}\{rudeness_{jt} > 0\}$	–0.01 (0.03)	0.03 (0.09)	0.00 (0.13)	–0.03 (0.12)	–0.03 (0.10)	–0.06 (0.09)
$\mathbb{1}\{dirtiness_{jt} > 0\}$	0.02 (0.03)	0.13 (0.06)	0.07 (0.12)	0.27 (0.12)*	0.04 (0.15)	–0.05 (0.09)
$congestion_{jt} \times \mathbb{1}\{rudeness_{jt} > 0\}$	–0.29 (0.20)	–0.51 (0.83)	–0.38 (0.95)	0.22 (0.75)	–0.29 (0.86)	0.16 (0.52)
$congestion_{jt} \times \mathbb{1}\{dirtiness_{jt} > 0\}$	–0.87 (0.24)***	–1.87 (0.52)***	–1.52 (1.18)	–2.11 (1.03)*	–1.52 (1.31).	–0.07 (0.52)
$congestion_{jt} \times notfree_j$	–0.26 (0.32)					
$congestion_{jt} \times outdoors_j$ ‡	2.98 (1.69).					
log odds($bias_i$)		1.00 (0.02)***	0.98 (0.02)***	0.96 (0.02)***	0.98 (0.01)***	1.00 (0.02)***
number of observations	375,345	44,602	14,259	12,896	26,875	38,941
number of regressors	122,615	43	43	43	43	
pseudo-R2 (McFadden)	0.7786	0.37801	0.38001	0.29541	0.36319	0.38208
AIC	48,077	14,388	15,481	28,759	45,737	

†Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

‡Only two museums are classified as outdoors museums.

from locals. The model in Table 10 includes the covariate $is.local_{i;j}$, which takes value one if the visitor at British (Prado) museum is from London (Madrid) and zero otherwise; it has also fixed effects for quarters. The data shows that, although local visitors' scores are higher than otherwise,

Table 10 British and Prado Museums. FEGLM Coefficient Estimates (standard errors in brackets).

	British M.	Prado
$congestion_{jt}$	–1.95 (0.58)***	–0.64 (0.40)
$is{:}local_{i;j}$	0.30 (0.11)**	1.46 (0.15)***
$congestion_{jt} \times local_i$	–2.73 (1.16)*	–3.38 (1.82).
log odds($bias_i$)	0.99 (0.02)***	0.98 (0.02)***
number of observations	34,755	20,992
of which, from locals	2,825	2,245
number of regressors	42	41
pseudo-R2 (McFadden)	0.13	0.14
AIC	37,310	22,008

†Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

they penalize congestion much more than non-locals, as per the coefficients of $congestion_{jt} \times local_i$ being negative, large in absolute value, and significant (p-values are .02 and .06 respectively).

6.2. Positive Quality Shocks

We close this section by inspecting the Prado Museum in Madrid, which houses numerous masterpieces by great universal artists. During the 10-year period considered in our analysis, it hosted more than 60 temporal exhibitions, of which the one devoted to Dutch painter Hieronymus Bosch—running from June 1 to September 25, 2016—was the most important, as measured by total number of visitors (almost 600,000), visitors per day (more than 5,000), caliber of the master pieces exhibited—including acclaimed The Garden of Earthly Delights—, press coverage, and critic reviews. We can thus interpret this exhibition as a positive quality shock to the museum content, during a four-month period. The setting is similar to a natural experiment, although there is no control group available given that the shock affected all visitors. It should thus be seen as an event study in which the treatment variable (visiting the museum during when the temporal exhibition was active) is expressed as $exhibition_t$.

We can assess the impact of the quality of the Bosch temporary exhibition on visitor sensitivity to congestion. We do so by using a model which includes the four covariates shown in Table 11 plus quarters as fixed effects in the year of the exhibition.

Table 11 Prado Museum. Impact of Bosch's temporal Exhibition. FEGLM Coe cient Estimates (standard errors in brackets).

$congestion_{jt}$	–3.94 (1.55)* †
$exhibition_t$	–0.04 (0.09)
$congestion_{jt}$ $exhibition_t$	3.63 (1.63)*
$\log \text{odds}(bias_i)$	0.99 (0.03)***
number of observations	5,828
number of regressors	8
pseudo-R2 (McFadden)	0.38
AIC	6,180

†Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

The model shows that Prado visitors clearly dislike congestion. However, during the Bosch exhibition, that sensitivity to congestion was substantially less negative, as per the absolute value of the coefficient of $congestion_{jt}$ $exhibition_t$ being large and positive. In fact, as Figure 10 shows, reported congestion during the Bosch's temporary exhibition is the highest of the times series. Yet, in contrast to any other busy periods at the museum, rating did not worsen in those four months.

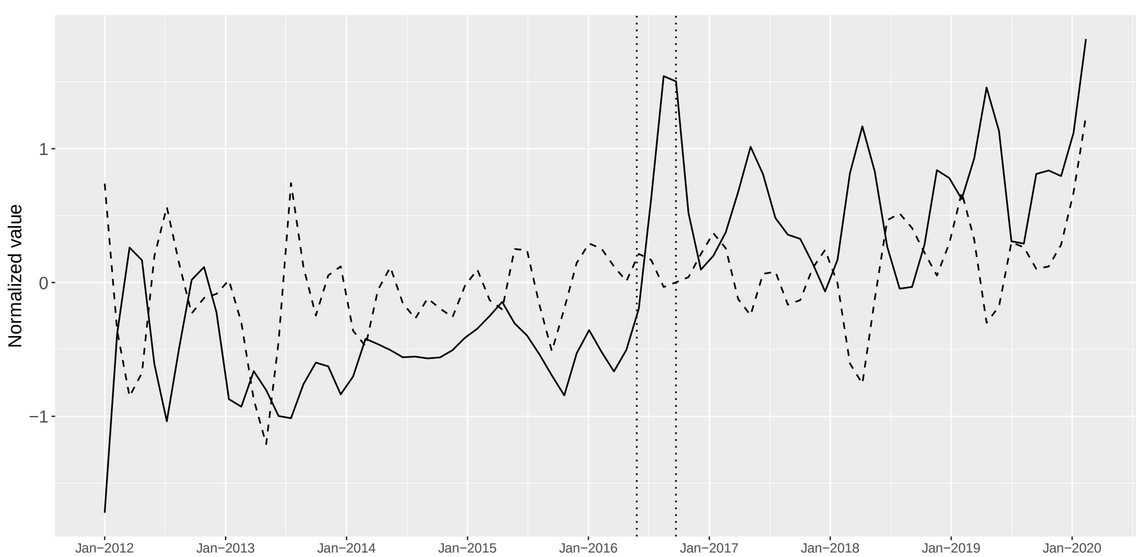

Figure 10 **Reported congestion (solid line) and visitors' utility (dashed line) over time. Note that, in contrast to subsequent periods with congestion peaks, utility did not decrease dramatically during the second semester of 2016, when Hieronymus Bosch's temporal exhibition was held.**

This finding allows us to conjecture – for lack of more comprehensive evidence – that significantly improving the inherent quality of an attraction can be used as a powerful tool to offset the impact of congestion on visitors' perception. Indeed, offering higher quality may be associated visitors expecting high congestion, which serves as a information mechanism to reduce sensitivity to waiting (Larson 1987). One could even think that quality can be used as an implicit channel to manage visitor expectations, as opposed to direct notification of expected congestion levels.

7. Conclusions and Managerial Insights

In this paper, we have developed a framework to study the impact of congestion on the satisfaction of visitors in touristic attractions, using public review data from multiple sites. We take advantage of the variation in service conditions (nature of the attraction, price, operating outcomes) to study the drivers that impact the sensitivity of satisfaction to congestion. Our approach allows us to control for visitor reporting biases and identify sensitivity from short-term (within the same quarter) variations in reported congestion.

Our results indicate that congestion clearly erodes satisfaction and that sensitivity to congestion does not depend on the quality of the attraction nor, by and large, the level of congestion. In contrast, it is heavily dependent on the nature of the attraction: zoos, aquariums and amusement parks are the most sensitive to congestion, while places of worship, city centers, and ancient ruins locations are the least sensitive. Beyond this main driver, we find that staff attentiveness and facilities cleanliness significantly mitigate the negative impact of congestion. Even though those two operational variables are not directly connected to congestion, visitors become less tolerant of queues when treated discourteously or exposed to uncleanliness. We also provide evidence that local visitors are more intolerant to congestion vis-à-vis tourists. Furthermore, we illustrate how positive quality shocks can reduce the negative impact of congestion.

Our findings offer a number of insights for managers in the service arena. First, they can map the sensitivity to congestion of the attraction they manage, in a chart linking dissatisfaction-prevalence together with other attractions within the same type (as in Appendix F), for benchmark purposes. The mapping can unveil the need to improve in two dimensions: attractions at the bottom (e.g., $\hat{\beta}_j$ 2) should watch congestion more carefully that those at the top. The particular actions to be taken beyond the obvious ones— admit fewer visitors—are indeed attraction-dependent.

Second, given that staff attentiveness and cleanliness impact satisfaction regarding congestion, potential actionable items include training staff to be more considerate with visitors and reinforce cleaning processes, especially on busy days. These actions can have a double impact on rating: a trivial one affecting satisfaction directly, and an indirect one through a better tolerance of congestion.

Third, given that positive quality shocks may reduce sensitivity to congestion, one other improvement point is to increase the inherent quality of attraction, especially in busy periods. This will lead to more patient visitors, as the attraction will be worth the waiting. Street bands playing cheerful music or cartoon characters being photographed with visitors at amusement parks on busy days are superb examples of this strategy. Also, given that locals are more intolerant to congestion than tourists, the former should be nudged to visit attractions on uncrowded days, thereby reducing congestion and increasing satisfaction also for those who visit on busy days.

Finally, our work can be expanded in various directions. Generally, our methodology can be readily extended to measure sensitivity of satisfaction to operational outcomes in service settings, provided that there is a corpus of reviews from which to identify the outcome of interest. Here we have chosen congestion as the main variable of analysis, but other variables may be the focus of similar studies, such as price, organizational choices, or staff behavior. Other service contexts different from attractions may be suitable to conduct similar research, including hotels, restaurants or hospitals. The first two contexts may be immediate applications given that there is plenty of data accessible from various sources (e.g., Airbnb, Yelp, Booking.com, Trip Advisor). When doing so, the main variable of study should depend on the category—e.g., staff politeness at hotels, quality of food at restaurants—as well as its drivers—e.g., cleanliness and noise at hotels, staff politeness and waiting time at restaurants.

Acknowledgments

Victor Martínez-de-Albéniz acknowledges the financial support provided by the Agencia Estatal de Investigación (AEI) from the Spanish Ministry of Science and Innovation, with project reference PID2020-116135GB-I00 MCIN/ AEI / 10.13039/501100011033.

References

Allon G, Federgruen A, Pierson M (2011) How much is a reduction of your customers' wait worth? An empirical study of the fast-food drive-thru industry based on structural estimation methods. *Manufacturing & Service Operations Management* 13(4):489–507.

Allon G, Kremer M (2018) Behavioral foundations of queueing systems. *The Handbook of Behavioral Operations* 9325:325–366.

Allon G, Zhang D (2017) Managing service systems in the presence of social networks, working paper, available at SSRN 2673137.

Antonides G, Verhoef PC, Van Aalst M (2000) Consumer perception and evaluation of waiting time: A field experiment. *Report Series. Research in Management. Erasmus Research Institute of Management* .

Ba S, Pavlou PA (2002) Evidence of the effect of trust building technology in electronic markets: Price premiums and buyer behavior. *MIS quarterly* 243–268.

Baker J, Cameron M (1996) The effects of the service environment on affect and consumer perception of waiting time: An integrative review and research propositions. *Journal of the Academy of marketing Science* 24(4):338–349.

Berge L (2018) Efficient estimation of maximum likelihood models with multiple fixed-effects: the R package FENmlm. *CREA Discussion Papers, 13* .

Bitran GR, Ferrer JC, Rocha e Oliveira P (2008) OM forum—managing customer experiences: perspectives on the temporal aspects of service encounters. *Manufacturing & Service Operations Management* 10(1):61–83.

Chen N, Li A, Talluri K (2021) Reviews and self-selection bias with operational implications. *Management Science* 67(12):7472–7492.

Chen Z, Lurie NH (2013) Temporal contiguity and negativity bias in the impact of online word of mouth. *Journal of Marketing Research* 50(4):463–476.

Chevalier JA, Mayzlin D (2006) The effect of word of mouth on sales: Online book reviews. *Journal of marketing research* 43(3):345–354.

Cox DR, Smith WL (1961) *Queues* (Chapman and Hall/CRC).

Dellarocas C (2006) Strategic manipulation of internet opinion forums: Implications for consumers and firms. *Management science* 52(10):1577–1593.

Dellarocas C, Gao G, Narayan R (2010) Are consumers more likely to contribute online reviews for hit or niche products? *Journal of Management Information Systems* 27(2):127–158.

Deshmane A, Martínez-de-Albéniz V, Roels G (2023) Intertemporal spillovers in consumer experiences: Empirical evidence and service design implications, working paper, IESE Business School.

Hollenbeck B, Moorthy S, Proserpio D (2019) Advertising strategy in the presence of reviews: An empirical analysis. *Marketing Science* 38(5):793–811.

Houston MB, Bettencourt LA, Wenger S (1998) The relationship between waiting in a service queue and evaluations of service quality: A field theory perspective. *Psychology & Marketing* 15(8):735–753.

Hu K, Xu X, Qu A (2021) The psychology of virtual queue: When waiting feels less like waiting, working paper, available at SSRN 3933040.

Huang N, Burtch G, Hong Y, Polman E (2016) Effects of multiple psychological distances on construal and consumer evaluation: A field study of online reviews. *Journal of Consumer Psychology* 26(4):474–482.

Katz K, Larson R, Larson B (1991) Prescription for waiting–in line blues: Entertain, enlighten and engage. *Sloan Management Review* 32(2):44–55.

Koo M, Fishbach A (2010) A silver lining of standing in line: Queuing increases value of products. *Journal of Marketing Research* 47(4):713–724.

Kremer M, Debo L (2016) Inferring quality from wait time. *Management Science* 62(10):3023–3038.

Larson RC (1987) Or forum—perspectives on queues: Social justice and the psychology of queueing. *Operations research* 35(6):895–905.

Lu Y, Musalem A, Olivares M, Schilkrut A (2013) Measuring the effect of queues on customer purchases. *Management Science* 59(8):1743–1763.

Maister DH (1985) *The psychology of waiting lines* (Harvard Business School).

Martínez-de-Albéniz V, Sosa M, Carrera C (2020) The bright side of bad choices: Evidence from restaurant exploration, working paper, IESE Business School.

Mejía J, Mankad S, Gopal A (2021) Service quality using text mining: Measurement and consequences. *Manufacturing & Service Operations Management* 23(6):1354–1372.

Parasuraman A, Zeithaml VA, Berry LL (1985) A conceptual model of service quality and its implications for future research. *Journal of marketing* 49(4):41–50.

Perdikaki O, Kesavan S, Swaminathan JM (2012) Effect of traffic on sales and conversion rates of retail stores. *Manufacturing & Service Operations Management* 14(1):145–162.

Rafaeli A, Barron G, Haber K (2002) The effects of queue structure on attitudes. *Journal of service research* 5(2):125–139.

Smith WE (1955) Various optimizers for single-stage production. Technical report, California Univ. Los Angeles Numerical Analysis Research.

Sulek JM, Hensley RL (2004) The relative importance of food, atmosphere, and fairness of wait: The case of a full-service restaurant. *Cornell Hotel and Restaurant Administration Quarterly* 45(3):235–247.

Taylor S (1994) Waiting for service: the relationship between delays and evaluations of service. *Journal of marketing* 58(2):56–69.

Ülkü S, Hydock C, Cui S (2020) Making the wait worthwhile: Experiments on the effect of queueing on consumption. *Management Science* 66(3):1149–1171.

Ülkü S, Hydock C, Cui S (2022) Social queues (cues): Impact of others' waiting in line on one's service time. *Management Science* 68(11):7793–8514.

Van Mieghem JA (1995) Dynamic scheduling with convex delay costs: The generalized c-mu rule. *The Annals of Applied Probability* 809–833.

Voorhees CM, Baker J, Bourdeau BL, Brocato ED, Cronin Jr JJ (2009) It depends: Moderating the relationships among perceived waiting time, anger, and regret. *Journal of Service Research* 12(2):138–155.

Zhou R, Soman D (2008) Consumers' waiting in queues: The role of first-order and second-order justice. *Psychology & Marketing* 25(3):262–279.

Appendices

A. Output of regression models of Section 5.7 (Robustness)

Table 12 Top 500 attractions. Coefficient estimates in simple models (standard errors in brackets).

Model	0	0'	1	1'	2	3
Covariates						
$congestion_{jt}$		–1.19 (0.00)*** †		–1.08 (0.03)***		–0.82 (0.03)***
Fixed effects						
User	No	No	Yes	Yes	Yes	Yes
Attraction	Yes	Yes	Yes	Yes	No	No
Attraction-quarter	No	No	No	No	Yes	Yes
number of observations	8,768,806	8,768,806	4,741,612	4,741,612	4,741,133	4,741,133
number of regressors	500	501	848,301	848,302	865,715	865,716
AIC	10,798,309	10,791,502	6,880,373	6,878,462	6,854,513	6,853,732

†Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table 13 Top 500 attractions. Coefficient estimates in models 3 to 5 (standard errors in brackets) whit alternative definition of excellent satisfaction.

Model	3	4	5
Covariates			
$congestion_{jt}$	–1.02 (0.05)*** †		
$congestion_{jt}$ $amusementpark$		–1.80 (0.15)***	–1.58 (0.24)***
$congestion_{jt}$ $ancientruins$		–0.29 (0.35)	–0.12 (0.49)
$congestion_{jt}$ $aquarium$		–2.80 (0.46)***	–2.72 (0.49)***
$congestion_{jt}$ $building$		–1.13 (0.15)***	–1.09 (0.20)***
$congestion_{jt}$ $cityarea$		–0.58 (0.09)***	–0.33 (0.18).
$congestion_{jt}$ $museum$		–1.22 (0.12)***	–1.20 (0.17)***
$congestion_{jt}$ $naturalfeature$		–0.85 (0.17)***	–0.61 (0.25)***
$congestion_{jt}$ $palace/castle$		–1.34 (0.18)***	–1.34 (0.23)***
$congestion_{jt}$ $placeofworship$		–0.67 (0.17)***	–0.71 (0.21)***
$congestion_{jt}$ $shoppingarea$		–0.58 (0.26)*	–0.48 (0.26).
$congestion_{jt}$ zoo		–3.56 (0.63)***	–3.45 (0.65)***
$\mathbf{1}\{rudeness_{jt} > 0\}$ ‡			–0.08 (0.01)***
$\mathbf{1}\{dirtiness_{jt} > 0\}$			–0.10 (0.01)***
$congestion_{jt}$ $\mathbf{1}\{rudeness_{jt} > 0\}$			–0.26 (0.10)**
$congestion_{jt}$ $\mathbf{1}\{dirtiness_{jt} > 0\}$			–0.17 (0.08)*
$congestion_{jt}$ $notfree_j$			0.16 (0.15)
$congestion_{jt}$ $outdoors_j$			–0.23 (0.16)
Fixed effects			
User	Yes	Yes	Yes
Attraction	No	No	No
Attraction-quarter	Yes	Yes	Yes
number of observations	2,715,506	2,715,506	2,715,506
number of regressors	441,666	441,676	441,682
AIC	3,067,661	3,067,549	3,066,838

†Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

‡$\mathbf{1}\{\}$ represents the indicator function.

Table 14 Top 500 attractions. Coefficient estimates in models 3 to 5 (standard errors in brackets) using the probit function.

Model	3	4	5
Covariates			
$congestion_{jt}$	–0.48 (0.02)*** †		
$congestion_{jt}$ $amusementpark$		–0.80 (0.06)***	–0.72 (0.10)***
$congestion_{jt}$ $ancientruins$		–0.29 (0.12)*	–0.30 (0.14)*
$congestion_{jt}$ $aquarium$		–1.07 (0.19)***	–0.99 (0.19)***
$congestion_{jt}$ $building$		–0.56 (0.06)***	–0.52 (0.08)***
$congestion_{jt}$ $cityarea$		–0.30 (0.04)***	–0.34 (0.07)***
$congestion_{jt}$ $museum$		–0.56 (0.05)***	–0.49 (0.06)***
$congestion_{jt}$ $naturalfeature$		–0.53 (0.07)***	–0.55 (0.10)***
$congestion_{jt}$ $palace{=}castle$		–0.53 (0.07)***	–0.46 (0.09)***
$congestion_{jt}$ $placeofworship$		–0.32 (0.06)***	–0.28 (0.07)***
$congestion_{jt}$ $shoppingarea$		–0.43 (0.11)***	–0.41 (0.11)***
$congestion_{jt}$ zoo		–1.44 (0.23)***	–1.41 (0.24)***
$\mathbf{1}\{rudeness_{jt} > 0\}$ ‡			–0.01 (.005)*
$\mathbf{1}\{dirtiness_{jt} > 0\}$			–0.02 (.004)***
$congestion_{jt}$ $\mathbf{1}\{rudeness_{jt} > 0\}$			–0.15 (0.04)***
$congestion_{jt}$ $\mathbf{1}\{dirtiness_{jt} > 0\}$			–0.17 (0.03)***
$congestion_{jt}$ $notfree_j$			–0.01 (0.05)
$congestion_{jt}$ $outdoors_j$			0.09 (0.07)
Fixed effects			
User	Yes	Yes	Yes
Attraction	No	No	No
Attraction-quarter	Yes	Yes	Yes
number of observations	4,741,133	4,741,133	4,741,133
number of regressors	865,716	865,726	865,732
AIC	6,855,826	6,855,725	6,855,235

†Significance codes: 0 ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’.’ 0.1 ’ ’ 1

‡$\mathbf{1}\{\}$ represents the indicator function.

Table 15 Top 500 attractions. Coefficient estimates in models 3 to 5 (standard errors in brackets) using a linear regression.

Model	3	4	5
Covariates			
$congestion_{jt}$	–0.22 (0.01)*** †		
$congestion_{jt}$ $amusementpark$		–0.48 (0.03)***	–0.39 (0.04)***
$congestion_{jt}$ $ancientruins$		–0.08 (0.04)*	–0.05 (0.05)
$congestion_{jt}$ $aquarium$		–0.53 (0.07)***	–0.49 (0.07)***
$congestion_{jt}$ $building$		–0.29 (0.02)***	–0.25 (0.03)***
$congestion_{jt}$ $cityarea$		–0.13 (0.01)***	–0.11 (0.03)***
$congestion_{jt}$ $museum$		–0.25 (0.02)***	–0.20 (0.02)***
$congestion_{jt}$ $naturalfeature$		–0.16 (0.02)***	–0.13 (0.04)***
$congestion_{jt}$ $palace{=}castle$		–0.28 (0.03)***	–0.23 (0.03)***
$congestion_{jt}$ $placeofworship$		–0.10 (0.02)***	–0.07 (0.03)**
$congestion_{jt}$ $shoppingarea$		–0.21 (0.04)***	–0.19 (0.04)***
$congestion_{jt}$ zoo		–0.58 (0.08)***	–0.53 (0.09)***
$\mathbf{1}\{rudeness_{jt} > 0\}$ ‡			–.005 (.002)*
$\mathbf{1}\{dirtiness_{jt} > 0\}$			–0.01 (.001)***
$congestion_{jt}$ $\mathbf{1}\{rudeness_{jt} > 0\}$			–0.13 (0.02)***
$congestion_{jt}$ $\mathbf{1}\{dirtiness_{jt} > 0\}$			–0.11 (0.01)***
$congestion_{jt}$ $notfree_j$			–.004 (0.02)
$congestion_{jt}$ $outdoors_j$			0.01 (0.03)
Fixed effects			
User	Yes	Yes	Yes
Attraction	No	No	No
Attraction-quarter	Yes	Yes	Yes
number of observations	8,768,806	8,768,806	8,768,806
number of regressors	3,606,378	3,606,388	3,606,394
AIC	20,047,581	20,047,100	20,045,159

†Significance codes: 0 ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’.’ 0.1 ’ ’ 1

‡$\mathbf{1}\{\}$ represents the indicator function.

Table 16 Top 500 attractions. Coefficient estimates in models 3 to 5 when current opinion is disregarded to compute $congestion_{jt}$ (standard errors in brackets).

	3	4	5
Covariates			
$congestion_{jt}$	–0.38 (0.03)*** †		
$congestion_{jt}$ $amusementpark$		–0.89 (0.11)***	–0.73(0.16)***
$congestion_{jt}$ $ancientruins$		–0.21 (0.20)	–0.18 (0.24)
$congestion_{jt}$ $aquarium$		–1.15 (0.31)***	–1.02 (0.33)**
$congestion_{jt}$ $building$		–0.46 (0.10)***	–0.38 (0.13)**
$congestion_{jt}$ $cityarea$		–0.12 (0.06)*	–0.11 (0.12)
$congestion_{jt}$ $museum$		–0.58 (0.08)***	–0.48 (0.11)***
$congestion_{jt}$ $naturalfeature$		–0.28 (0.11)**	–0.25 (0.16)
$congestion_{jt}$ $palace/castle$		–0.47 (0.11)***	–0.38 (0.15)*
$congestion_{jt}$ $placeofworship$		–0.09 (0.10)	–0.04 (0.12)
$congestion_{jt}$ $shoppingarea$		–0.30 (0.18).	–0.26 (0.18)
$congestion_{jt}$ zoo		–1.99 (0.38)***	–1.92 (0.40)***
$\mathbf{1}\{rudeness_{jt} > 0\}$ ‡			–0.03 (.009)**
$\mathbf{1}\{dirtiness_{jt} > 0\}$			–0.04 (.006)***
$congestion_{jt}$ $\mathbf{1}\{rudeness_{jt} > 0\}$			–0.24 (0.07)***
$congestion_{jt}$ $\mathbf{1}\{dirtiness_{jt} > 0\}$			–0.27 (0.05)***
$congestion_{jt}$ $notfree_j$			0.00 (0.09)
$congestion_{jt}$ $outdoors_j$			0.08 (0.11)
Fixed effects			
User	Yes	Yes	Yes
Attraction	No	No	No
Attraction-quarter	Yes	Yes	Yes
number of observations	4,741,133	4,741,133	4,741,133
number of regressors	865,716	865,726	865,732
AIC	6,854,336	6,854,243	6,853,732

†Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

‡$\mathbf{1}\{\}$ represents the indicator function.

B. Procedure to Detect Congestion in Comments

This appendix presents the procedure used to identified individual congestion in comments. After the procedure is applied, variable $b.cng$ will take value one if there is congestion detected and value zero otherwise. By default, no congestion is assumed.

First, letters are uncapitalized and accents—acute, grave, and circumflex—and other especial signs of punctuation—dieresis, cedillas, and tildes—removed in all six languages. Then, for each anchor word in the comment (see rows labeled "Anchor" in Table 17 for a list of words in each language), the following conditions are applied:

1. First, if there is a positive modifier ("Positive") within seven words from the anchor, then $b.cng = 0$ and the procedure moves to the next comment.
2. Then, if there are negative words ("Negative") within four words from the anchor, then $b.cng = 0$ and the procedure moves to the next comment.
3. Finally, if there is a modifier up to seven words before ("Before") or after ("After") the anchor, then $b.cng = 1$.

Table 17 Words used in the procedure to identify congestion.

English	
Anchor	queue, wait/s/ed/ing, crowd/s/ed, overcrowded, packed, line, row, column chain, throng, pack, cram, fill, congest, busy, people, bundersome, jammed overwhelming, visitors,
Positive	fast, quickly, agile, small, fluid, quick, efficiently
Before	lot, very, much, many, terrible, large, long, endure, bad, horrible, endless numerous, immense, unique, patience, delay, failure, time, exhausting oppressive, quite, handle, huge, longest, bit, busy, big, really, pretty, massive wait/s/ed/ing, unless, wrong, biggest, stood , slowly, crazy, hours
After	long, huge, extensive, terrible, eternal, sorrow, incredible, numerous considerable, endless, bad, time, delay, massive, regardless, hours
Negative	no, not
Spanish	
Anchor	fila, cola/s, espera/ar/ando/amos, saturado/a/os/as, concurrido/a, agobiante agobio, gente, agobiados/as, masificado
Positive	rapido/a, rapidisimo/a, rapidamente, agil, pequena
Before	mucho/a, bastante, demasiado, buena, muy, terrible, unico/a, malo, extensa horrible, pena, larga, aguantar, gran, interminable, numerosa, inmensa paciencia, esperar, fallo, tiempo, demora
After	larga, enorme, extensa, terrible, eterna, pena, increible/s, numerosa considerable, interminable, malo, esperar
Negative	no
Portuguese	
Anchor	linha, fila, espera, saturado/a, ocupado, oprimido, opressor, pessoas esperar, aguardar, demorar, malta, gente, esmagado/a, sobrecarregado cansativo, avassalador
Positive	rapido/a, agil, pequeno/a, fluido, rapidamente
Before	bom, muito/a/os/as, terrivel, extenso, horrivel, pena, longo, aguentar, otimo infinito, numeroso, imenso/a, ma, mau, ruim, esperar, unico, paciencia, falha tempo, atraso, demorar, cheio, horas, espera
After	longo, enorme, extenso/a, terrivel, eterno, tristeza, incrivel, numeroso consideravel, fim, ruim, espera, tempo, imenso/a, comprido/a, larga, mau ma, grande, espantoso, interminavel, horas
Negative	nao

For instance, both “There was not a long queue at the entrance.” and “There was a queue at the entrance.” result in $b.cng = 0$.

Table 17 (cont.) **Words used in the procedure to identify congestion.**

Italian	
Anchor	fila, coda/e, attesa, saturo, occupato/a, sopraffatto, travolgente, popolo aspetto/are/iamo/ato, affollato/a, persone, gente
Positive	veloce, agile, piccolo/a, fluido/a, velocissimo/a, poca, rapida
Before	troppo, buono/a, molto/a/issimo/issima, terribile, esteso/a, cattivo/a, orribile scusa, lungo/a, insopportabile, grande, infinito/a, numeroso/a, immenso/a aspetta, unico/a, pazienza, fallimento, ricato, ore
After	lungo/a, enorme, esteso/a, terribile, eterno/a, dolore, incredibile, numeroso/a notevole, infinito/a, cattivo/a, aspetta, tempo, immenso/a, troppo molto/a/issimo/issima, ore
Negative	non
French	
Anchor	ligne, file, attente, sature, occupe, deborde, personnes, attends/du/dre/dons monde, gens, queue
Positive	rapide, petit, agile, rapidement, fluide
Before	beaucoup, trop, bien, tres, terrible, extensif, horrible, desole, long/gue, endurer super, interminable, nombreux, immense, mauvais, attendre, unique, patience echec, temporisation, grande, evitez/er, surpeuple, sature, encombre, heures
After	longue, enorme, extensif, terrible, eternel, chagrin, incroyable, nombreux considerable, interminable, mauvais, attendre, temps, trop, grande, longtemps importante, surpeuple, sature, encombre, tout, heures
Negative	non, ne
German	
Anchor	warteschlange, warten, gewartet, wartet, wartete, uberfullt, menge, gedraenge beschaftigt, schlange, reihe, packen, gepackt, voll, menschenmasse/en, menschen besucher, uberwaeltigend, verzogerung, zeit, linie
Positive	schnell/e, klein/e, flussig/e, effizient/e
Before	viele, sehr, schreckliche, grosse, lange/en/es, schlechte, endlose, zahlreiche, enorm stunden, ganzen, uber, einige, ewigkeit
After	lang, riesig, umfangreich, schrecklich, ewig, trauer, unglaublich, zahlreich betraechtlich, endlos, schlecht, zeit, geduld, gereicht
Negative	nicht, ohne

.

C. Procedure to Detect Rudeness in Comments

The procedure used to obtain a binary variable for rudeness is identical to the one described in Appendix B. The corresponding key words are in Table 18.

Table 18 Words used in the procedure to identify rudeness.

English	
Anchor	staff, worker/s, employee/s, reception, personnel, service/s, manager/s attendant, recepcionist, treatment, coordinator, guide
Positive	kind, attentive, good, great, best, unequalled, charming, polite, helpful friendly
Before and After	unpleasant, unfriendly, disrespectful, appalling, atrociuous, rude/ness, bad/ly terrible, awful, miserable, hostile, unhelpful, surly, aggressive, cynical poor, unkind, pissed, angry, disgusting, nasty, annoying, tempered, clumsy awkward, lazycrazy, dismissive, impolite, arrogant, useless, ignored nasty, disappointing, rigrid, horrific
Negative	no, not
Spanish	
Anchor	personal, trabajador/a/as/es, empleado/a/os, staff, trabaja/n, recepcion encargado/a/s/as, trato, trataron, guia
Positive	amable/s, atento/s, bueno/a, genial, mejor, inigualable, encantador/es
Before and After	decepcionante, mal/o/a/isimo, pesimo/a, terrible, grosero/s, desagradable/s mediocre/s, antipatico/s, maleducado/s, borde, agresivo/s, irrespetuoso, fatal desastre, decepcionados, lamentable, nefasto
Negative	no
Portuguese	
Anchor	pessoal, trabalhador/es, empregado/s, trabalho, tratamento, recepcao recepcionista, staff, manager, responsavel, gerente, guia, cansativo, avassalador
Positive	bondoso, atencioso, bom, atento, melhor, inigualavel, encantador incomparavel, amigavel
Before and After	desagradaveis/el, rude/s, ofensivos, pessimo, deficiente, maltratados grosseiro/s, antipatico/s, grosseiro, mal-educados, hostil, horrivel mal, mau, negativamente
Negative	nao
Italian	
Anchor	personale, lavoratori/e, impiegato/i, staff, servizio, lavoranti/e, reception trattati, pubblico, dipendenti/e, collaboratori, manager, responsabile direttore
Positive	gentile, attento, buonno, grande, migliore, ineguagliabile, affascinante cordiale, amichevoli, cortesi, simpatici
Before and After	maleducato/i, arrogante, sgradevole, irrispettoso, miserabile, aggressivo scortesi/e, antipatici, incapaci, mal/e/issimo, malamente, orribile terribile, pessimo
Negative	non

Table 18 (cont.) **Words used in the procedure to identify rudeness.**

French	
Anchor	employe/s, travailleur/s, salarie/e/s/es, travaillant, traites, traitement ouvrier, personnel, recepcion, employee, staff, manager/s, guide
Positive	gentil/s, amaible, sympathique, attentif/s, bon, bien
Before and After	desagreable/ment, nonchalant, impoli/e/s, hostile, irrespectueux, mauvais irrespectueuse, miserable, hargneux, cynique, deplaisants, horrible maltraites, grossier/s, rude/s, malpoli/s, peu agreable/s colere, mal
Negative	non, ne
German	
Anchor	arbeitnehmer, arbeiter, angestellter, personal, rezepcion, mitarbeiter behandelt, beschaftigten, fahrer,handwerker, staff, manager, direktor guide, besucher, uberwaeltigend, verzogerung, zeit, linie
Positive	freundlich/e, nett, aufmerksam, gut, besten, unvergleichlich, charmant
Before and After	unfreundlich, kantig, unhoflich, aggressiv, respektlos, murrisch, zynisch unangenehm, verletzend, schlecht, schrecklich/e, bose, furchtbar
Negative	nicht, ohne

D. Procedure to Detect Dirtiness in Comments

The binary variable that identifies dirtiness, *b.dirt*, takes value 1 whenever one or more key words in Table 18 (cont.) appear in the text, and zero otherwise.

Table 20 **Words used in the procedure to identify dirtiness.**

English	dirty, rubbish, filth, trash, refuse, dirtiness, unclean
Spanish	sucio/a/os/as, basura, porqueria, desperdicios, residuos
Portuguese	sujo/a, lixo, sujeira, porcaria, residuos
Italian	sporco/a, spazzatura, sporcizia, rifiuti, dirty, immondizia
French	ordures, dechets , sale/s, dirty, detritus
German	schmutzig, dreckig, mull, abfall

E. Selection of Rho

To obtain an adequate value of rho, ρ, we first assess the value of p, the (assumed constant) probability that a visitor at any attraction reports her visit on a given day provided that she has not reported it yet. In other words, we assume that the probability of reporting decreases exponentially with t: if T is the number of days elapsed between visit and report, $P\{T=r\}=p(1-p)^r, r=0,1,\dots$

Let N_s be the random number of visits on day s and X_t the random number of opinions on day t. Let $Z_{i;s;t}$, $s \leq t$, be a Bernoulli variable that takes value one if visitor i, who visits a given attraction on day s, reports on day t; and zero otherwise, that is $P\{Z_{i;s;t}=1\}=pq^{t-s}$, where $q=1-p$. The three random variables can be bound as follows.

$$X_t=\sum_{s\leq t}\sum_{i=1}^{N_s} Z_{i;s;t}$$

For a given attraction, we now derive the expected value of the number of opinions on day t, $\mathsf{E}(X_t)$, assuming that there is weekly "seasonality," that is, N_s has the same distribution as N_{s+7} for all s. Thus, we can write $\mathsf{E}(N_s)=\alpha_j\mu$, where μ is the average number of visits per week and α_j, $j=1,2,\dots,7$, are weights that only depend on the day of the week (here j is the day of the week where s belongs). Therefore,

$$\begin{aligned}\mathsf{E}(X_t)=&\mathsf{E}\left(\sum_{s\leq t}\sum_{i=1}^{N_s} Z_{i;s;t}\right)\\ =&\sum_{s\leq t}\mathsf{E}\left(\sum_{i=1}^{N_s} Z_{i;s;t}\right)\end{aligned}$$

Using the theorem of total expectation, it follows that

$$\begin{aligned}\mathsf{E}\left(\sum_{i=1}^{N_s} Z_{i;s;t}\right)=&\mathsf{E}_N\,\mathsf{E}_{Z|N}\left(\sum_{i=1}^{N_s} Z_{i;s;t}|N_s\right)\\ =&\mathsf{E}_N\left(\sum_{i=1}^{N_s}\mathsf{E}_{Z|N}Z_{i;s;t}\right)\\ =&\mathsf{E}_N\left(\sum_{i=1}^{N_s}P\{Z_{i;s;t}=1\}\right)\\ =&P\{Z_{i;s;t}=1\}\mathsf{E}_N\sum_{i=1}^{N_s}1\\ =&\mathsf{E}_N(N_s)P\{Z_{i;s;t}=1\}\end{aligned}$$

In particular, if $t=1$ (i.e., on Mondays),

$$\mathsf{E}(X_1)=\sum_{s\leq 1}\mathsf{E}(N_s)pq^{t-s}$$

$$
\begin{aligned}
&= \mathsf{E}(N_1)p + \mathsf{E}(N_0)pq + \dots \\
&= \alpha_1 \mu p + \alpha_7 \mu pq + \alpha_6 \mu pq^2 + \cdots + \alpha_1 \mu pq^7 + \alpha_7 \mu pq^8 + \dots \\
&= \mu p[\alpha_1(1 + q^7 + q^{14} + \dots) + \alpha_7 q(1 + q^7 + \dots) + \cdots + \alpha_2 q^6(1 + \dots)] \\
&= \frac{\mu p}{1 - q^7}(\alpha_1 + \alpha_7 q + \cdots + \alpha_2 q^6)
\end{aligned}
$$

If $t = 2$, likewise,

$$\mathsf{E}(X_2) = \frac{\mu p}{1 - q^7}(\alpha_2 + \alpha_1 q + \cdots + \alpha_3 q^6)$$

In general, for any value of t,

$$\mathsf{E}(X_t) = \frac{\mu p}{1 - q^7} \sum_{j=1}^{7} \alpha_{t-j+1} q^{j-1}, t = 1, 2, \dots, 7; \text{ and indeed } \mathsf{E}(X_t) = \mathsf{E}(X_{t-7}) \text{ if } t > 7.$$

where subscripts of α are congruent $(mod 7)$ to a number in $\{1, 2, \dots, 7\}$, e.g., $\alpha_{-2} = \alpha_5$.

The values of p and α_i are derived by solving the problem

$$\min_{p;\alpha_i} \sum_{i=1}^{7} \frac{(O_i - E_i)^2}{E_i}, \alpha_i \geq 0 \text{ for all } i, \sum_{i=1}^{7} \alpha_i = 1, \tag{2}$$

where O_i are the average number of opinions observed each day of the week and $E_i = \mathsf{E}(X_i)$.

The rationale for using this (chi-square) method is the following: The probability that a random review is written on day i of the week is $E(X_i)/\sum_{j=1}^{7} E(X_j) = E(X_i)/\mu$. We have a sample of $M\mu$ reviews, where M is the number of weeks. Then, the expected number of reviews on day i will be $ME(X_i)$. On the other hand, the number of observed reviews on day i is MO_i. Thus, we can see our problem as a goodness of fit one in a multinomial distribution with probabilities $E(X_i)/\mu$, $i = 1, \dots, 7$ and sample size $M\mu$. The chi-square statistic is

$$\sum_{i=1}^{7} \frac{(MO_i - ME_i)^2}{ME_i},$$

which is minimized for the same values as (2).

Note that since the number of bins is 7 and the number of (free) parameters is also 7, a perfect fit can be obtained in (2). However, it is interesting to have a model for the α's to find a more parsimonious estimation of p, which is our final goal. A good option is considering attractions that are closed during the weekdays, so $\alpha_1 = \alpha_2 = \cdots = \alpha_5 = 0$.

We consider two attractions that operate only in weekends: "Gangsters Tour," a cultural bus tour to the underworld in London, UK and "RSE Italy," a circuit in Imola, Italy where you can drive a Ferrari or a Lamborghini for fun. By doing so, we simplify the program just described, as α_i are zero if $i = 1, 2, \dots, 5$ thus $\alpha_6 = 1 - \alpha_7$. We use all observations of these attractions in our data set to obtain $O = \{0.46, 0.35, 0.13, 0.12, 0.17, 1.16, 1.10\}$ for Gangsters Tour and $O =$

$f1.87, 0.89, 0.56, 0.41, 0.35, 1.84, 2.84g$ for RSE Italy. Solving numerically for p we get $p = 0.371$ in the former attraction and $p = 0.379$ in the latter, which means that 96.3% of visitors will be expected to have submitted their reports to Trip Advisor six days after their respective visits (as $1 \quad (1 \quad 0.375)^{6+1} \quad 0.963$).

We then choose $\rho = 3$ so that the congestion average on day t over a week (i.e., $3 + 1 + 3$ days) includes roughly 95% of opinions from visitors at the attraction on day $t \quad \rho$ (and, admittedly, many more). In §5.7.5 we assess other values of ρ.

F. Congestion Charts for Individual Attraction Types

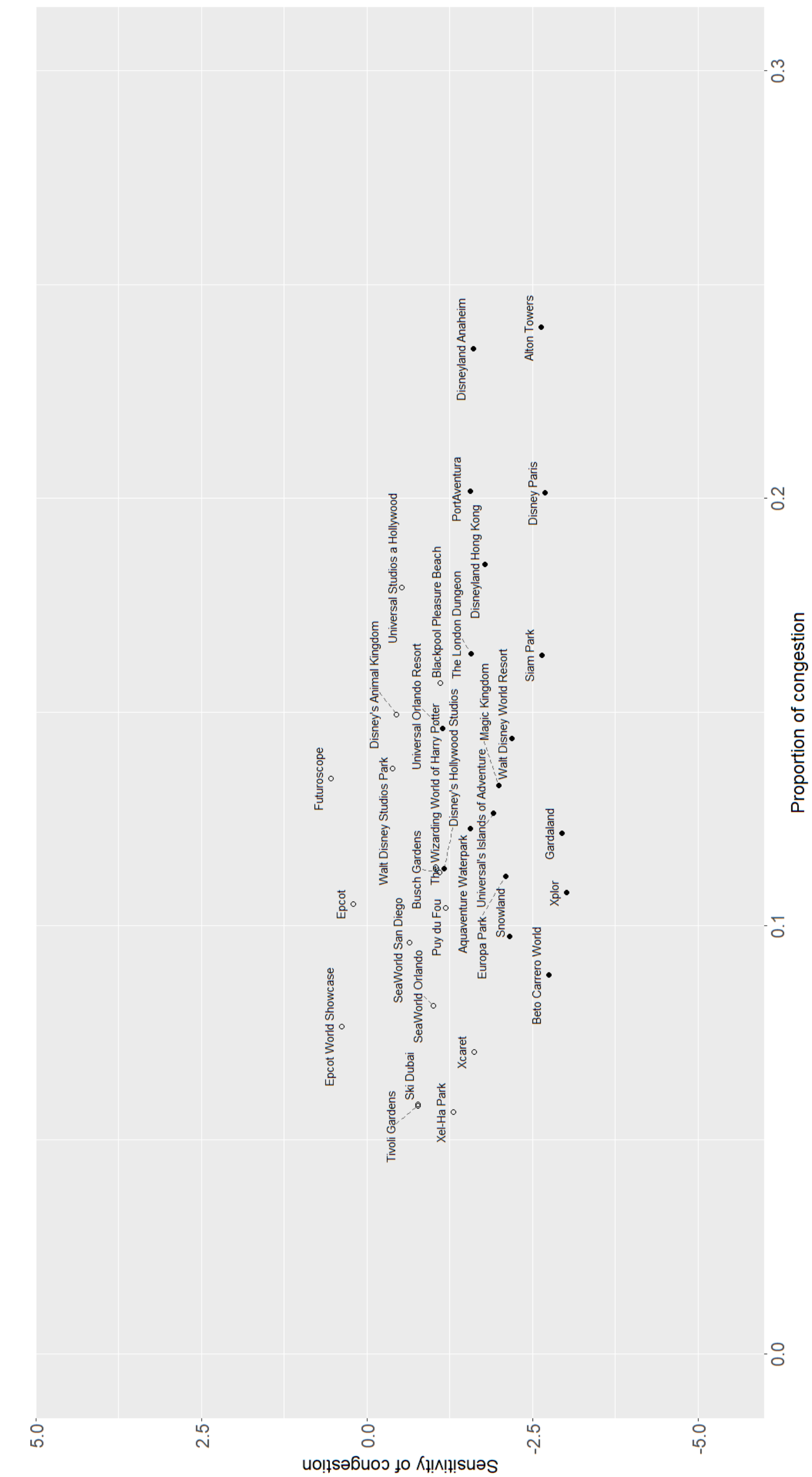

Figure 11 Top 500 Attractions. Congestion Chart for Amusement Parks (detail of Figure 5).

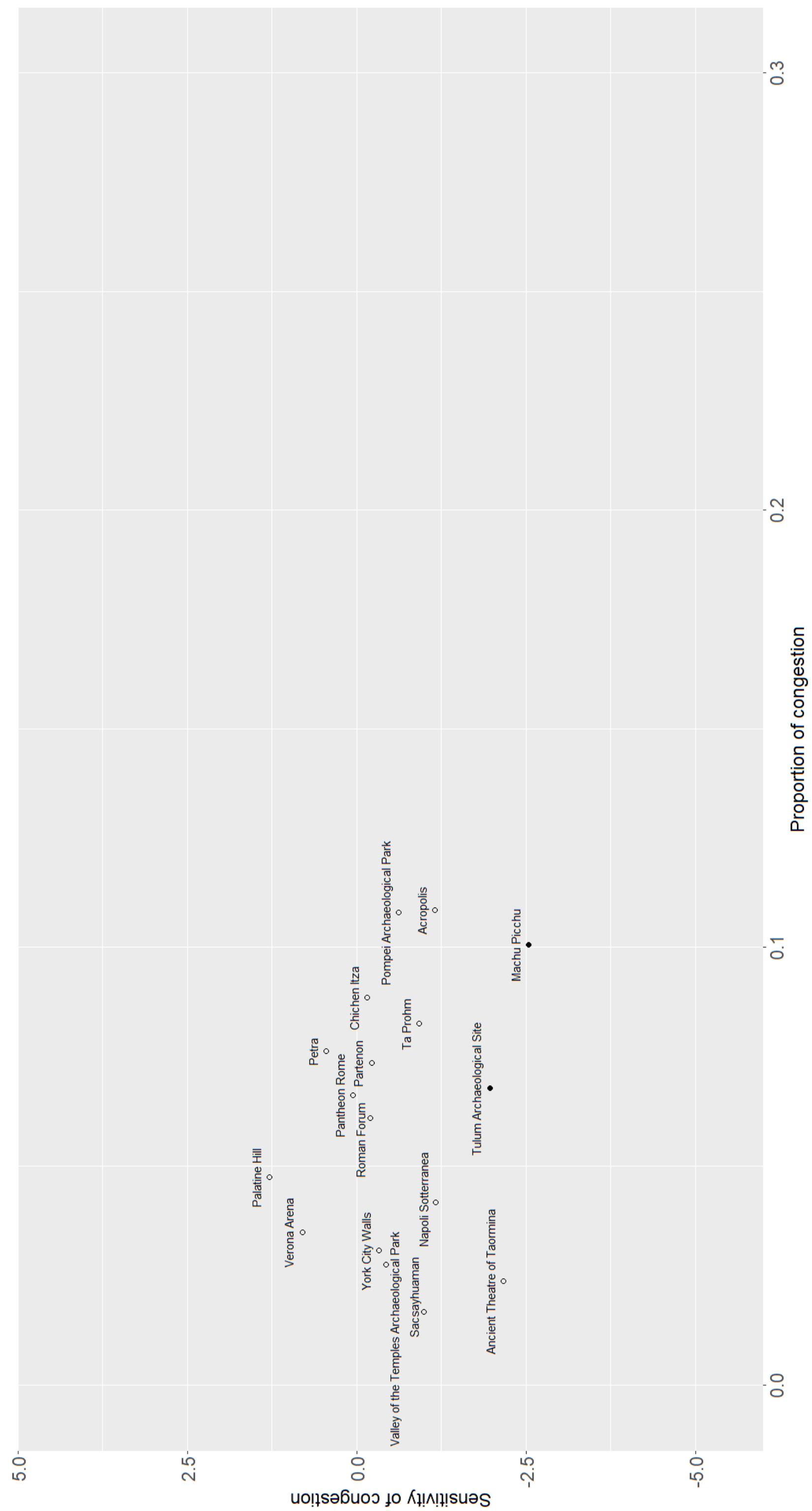

Figure 12 **Top 500 Attractions. Congestion Chart for Ancient Ruins (detail of Figure 5).**

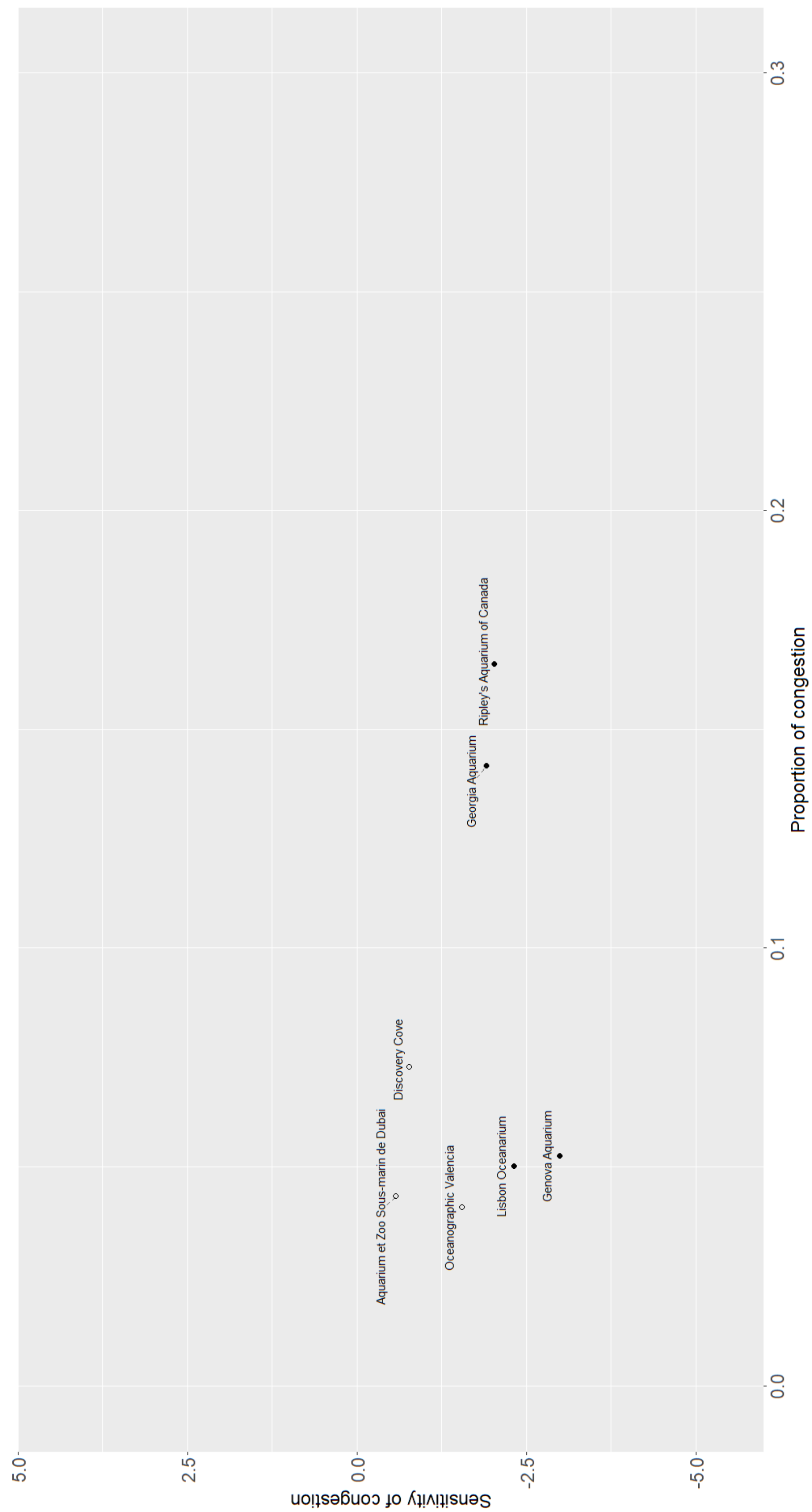

Figure 13 Top 500 Attractions. Congestion Chart for Aquariums (detail of Figure 5).

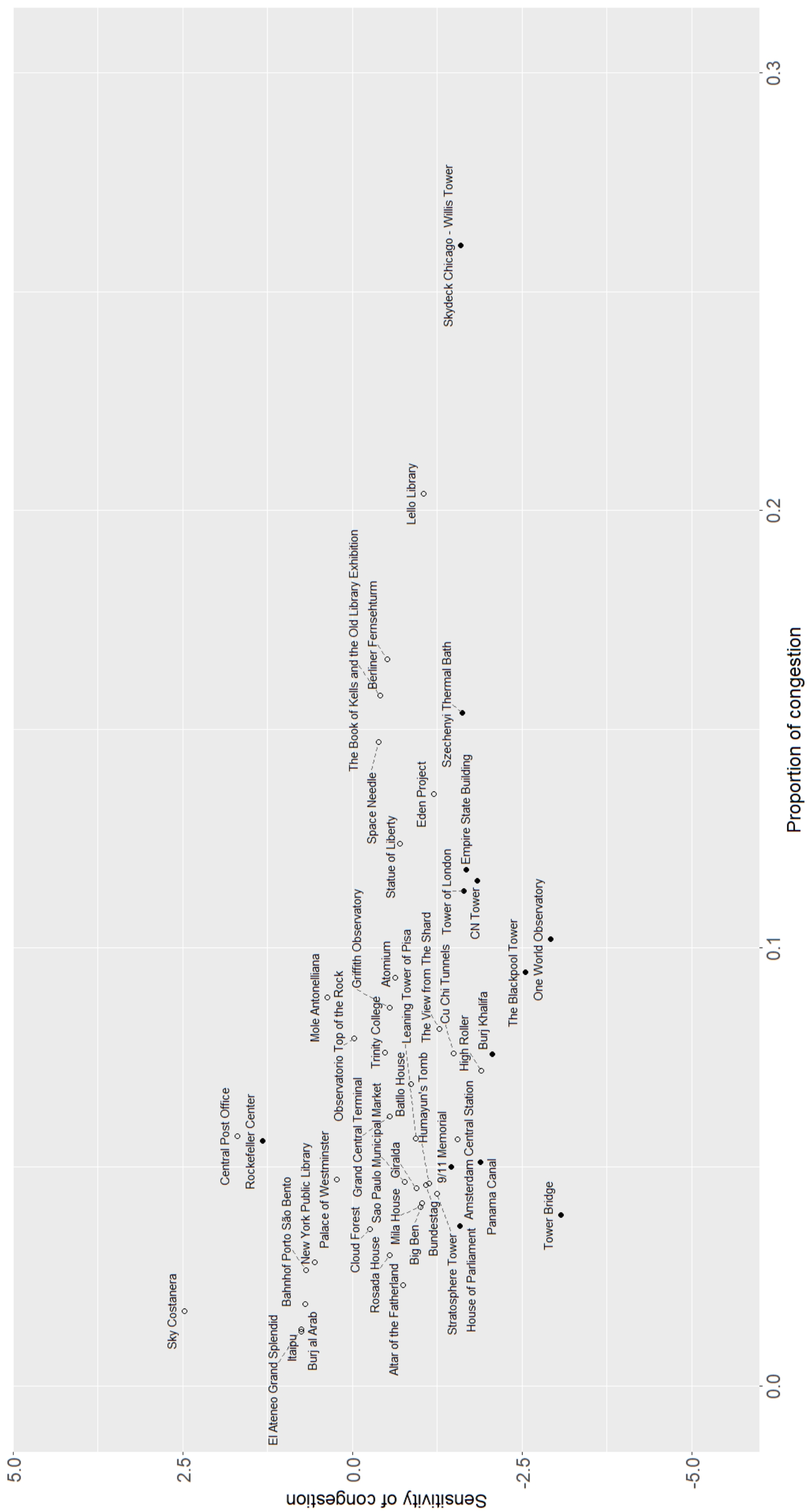

Figure 14 Top 500 Attractions. Congestion Chart for Buildings (detail of Figure 5).

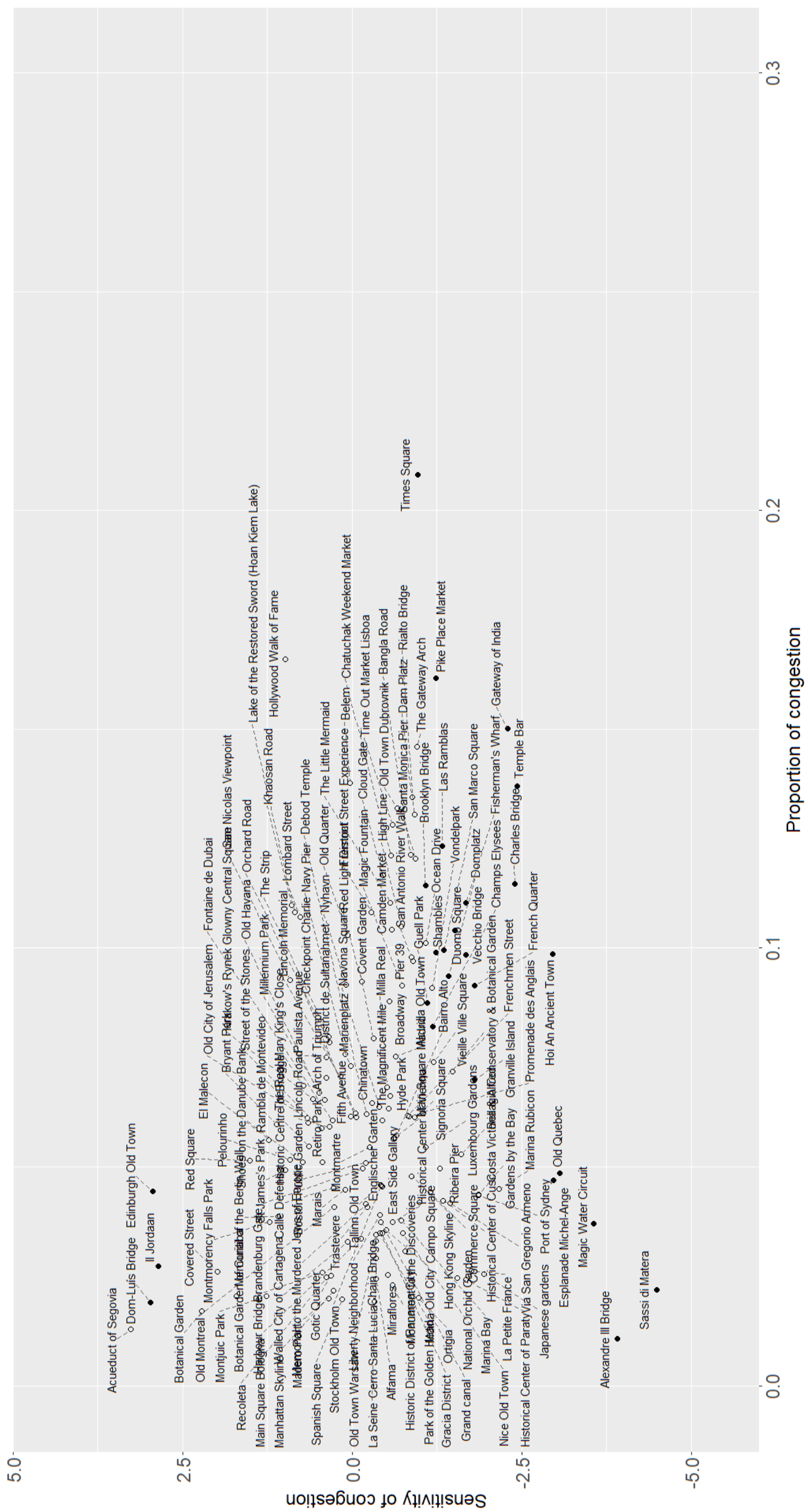

Figure 15 **Top 500 Attractions. Congestion Chart for City Areas (detail of Figure 5).**

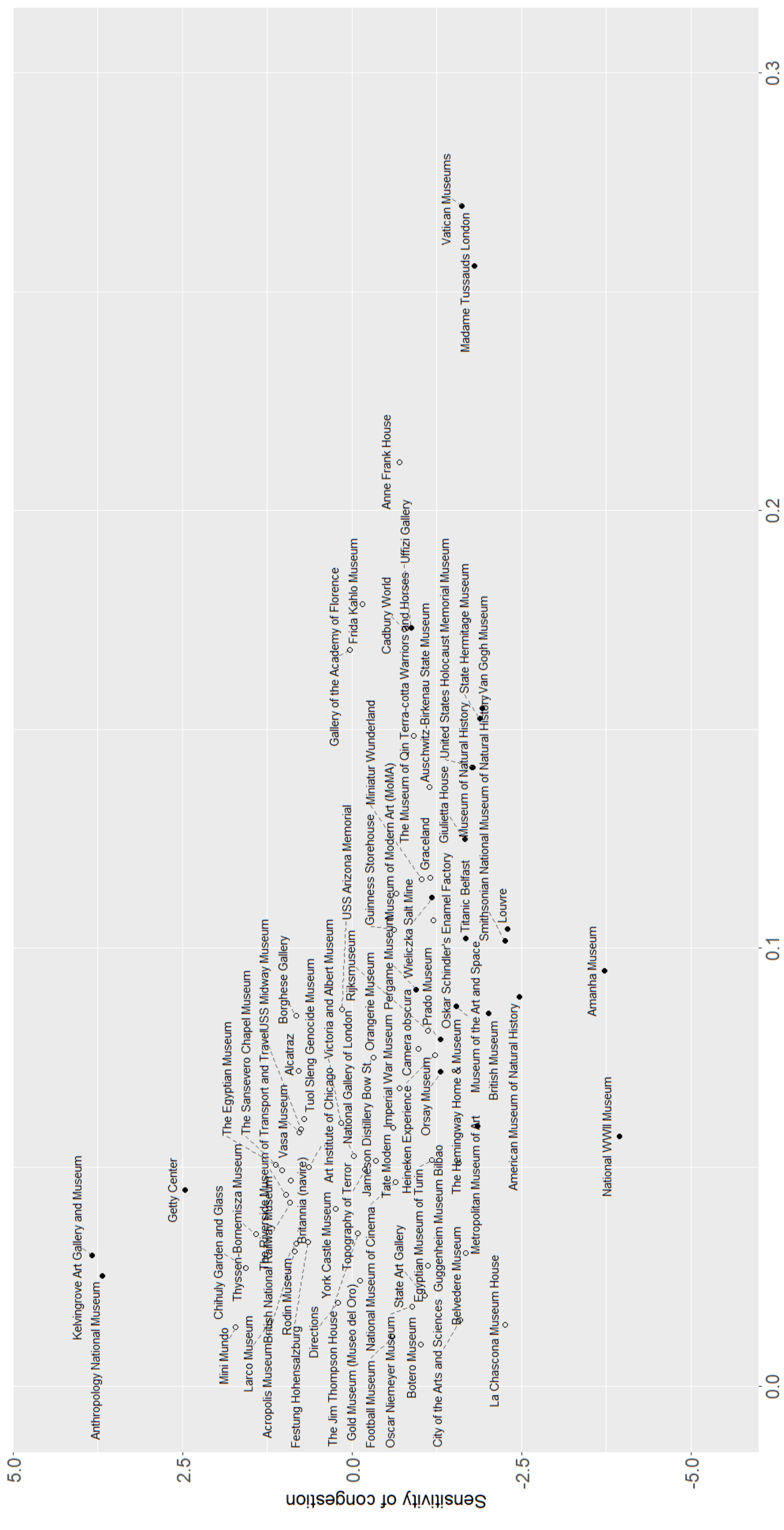

Figure 16 **Top 500 Attractions. Congestion Chart for Museums (detail of Figure 5).**

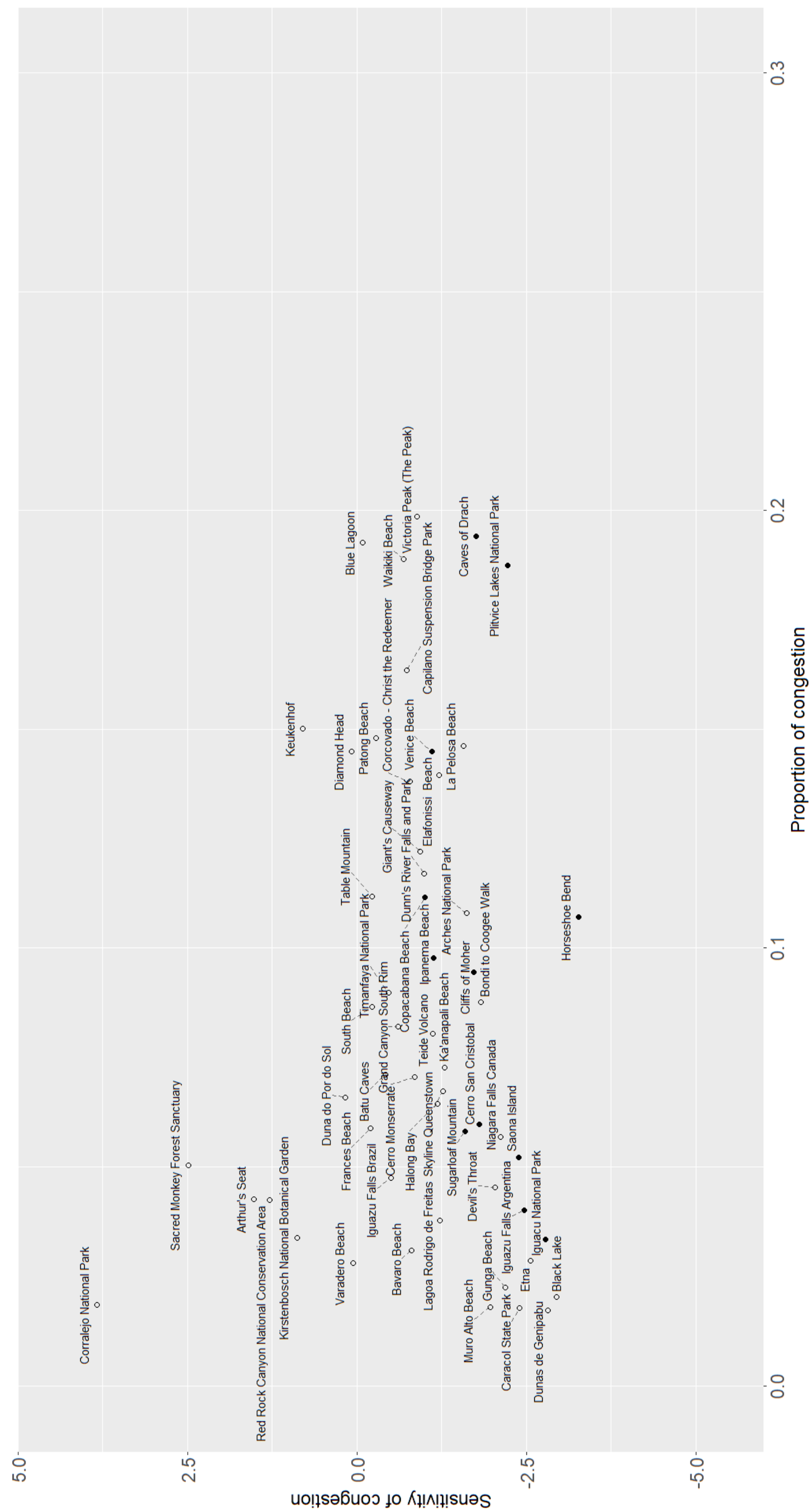

Figure 17 Top 500 Attractions. Congestion Chart for Natural Features (detail of Figure 5).

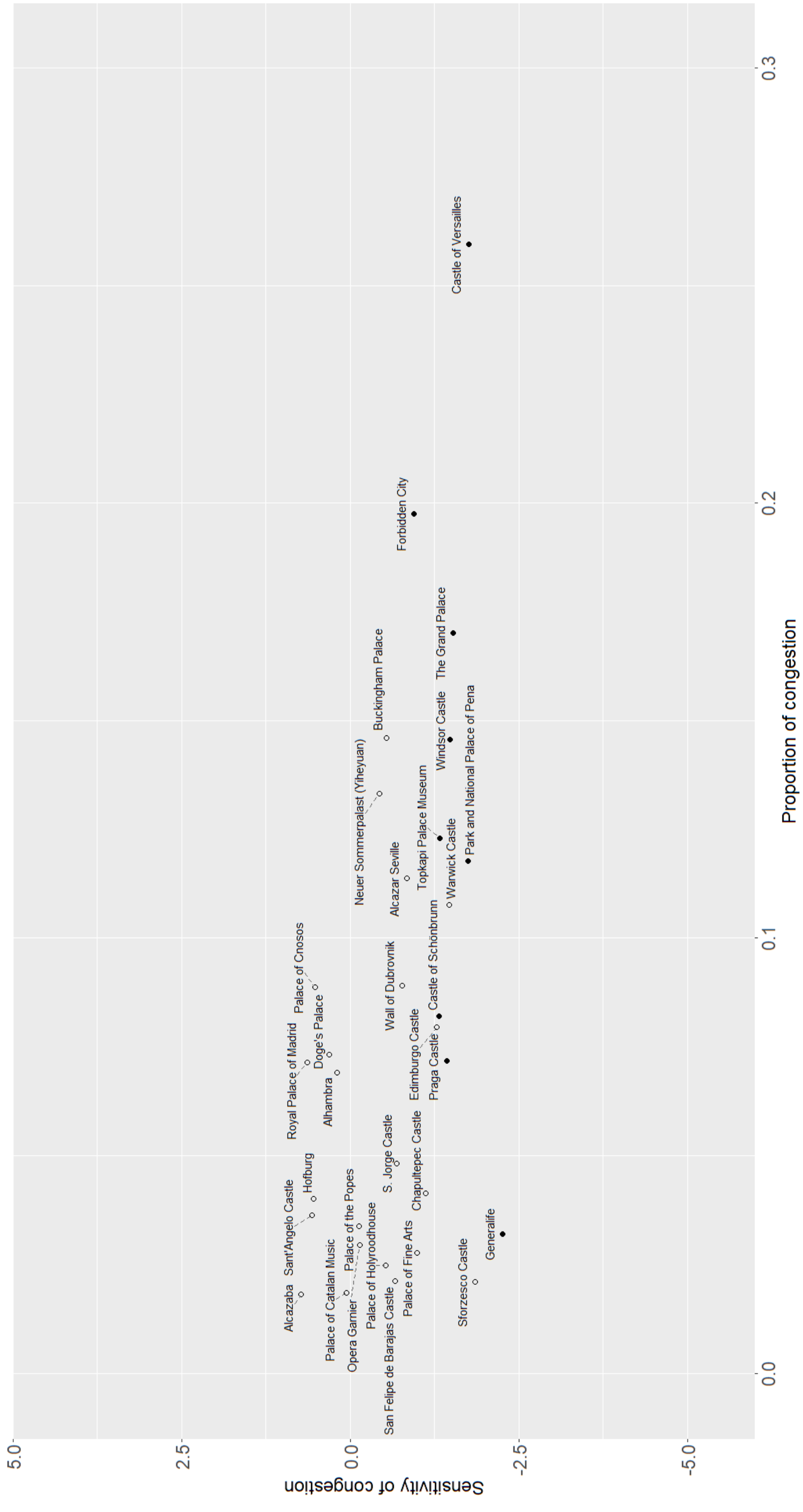

Figure 18 **Top 500 Attractions. Congestion Chart for Palaces/Castles (detail of Figure 5).**

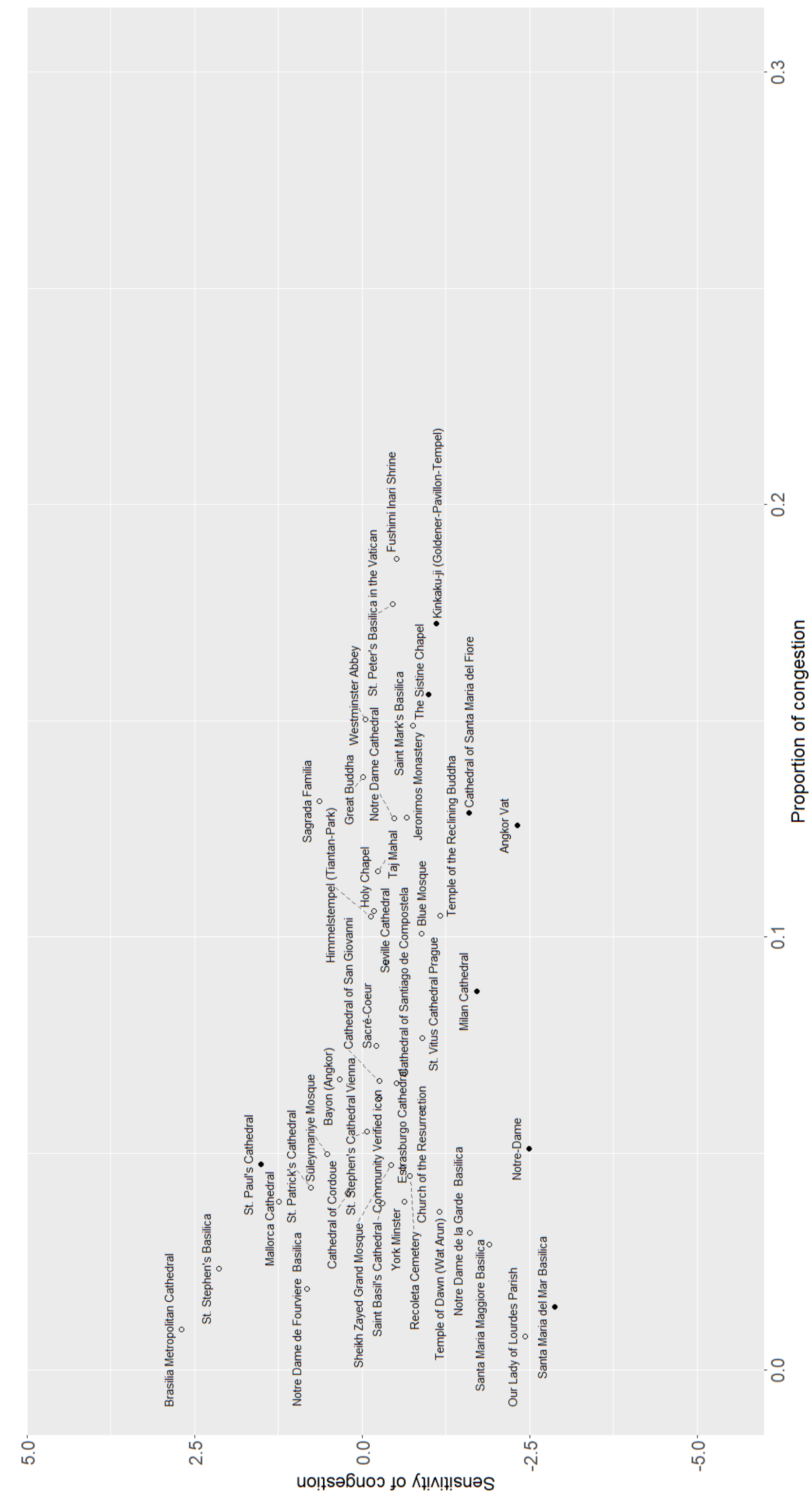

Figure 19 Top 500 Attractions. Congestion Chart for Places of Worship (detail of Figure 5).

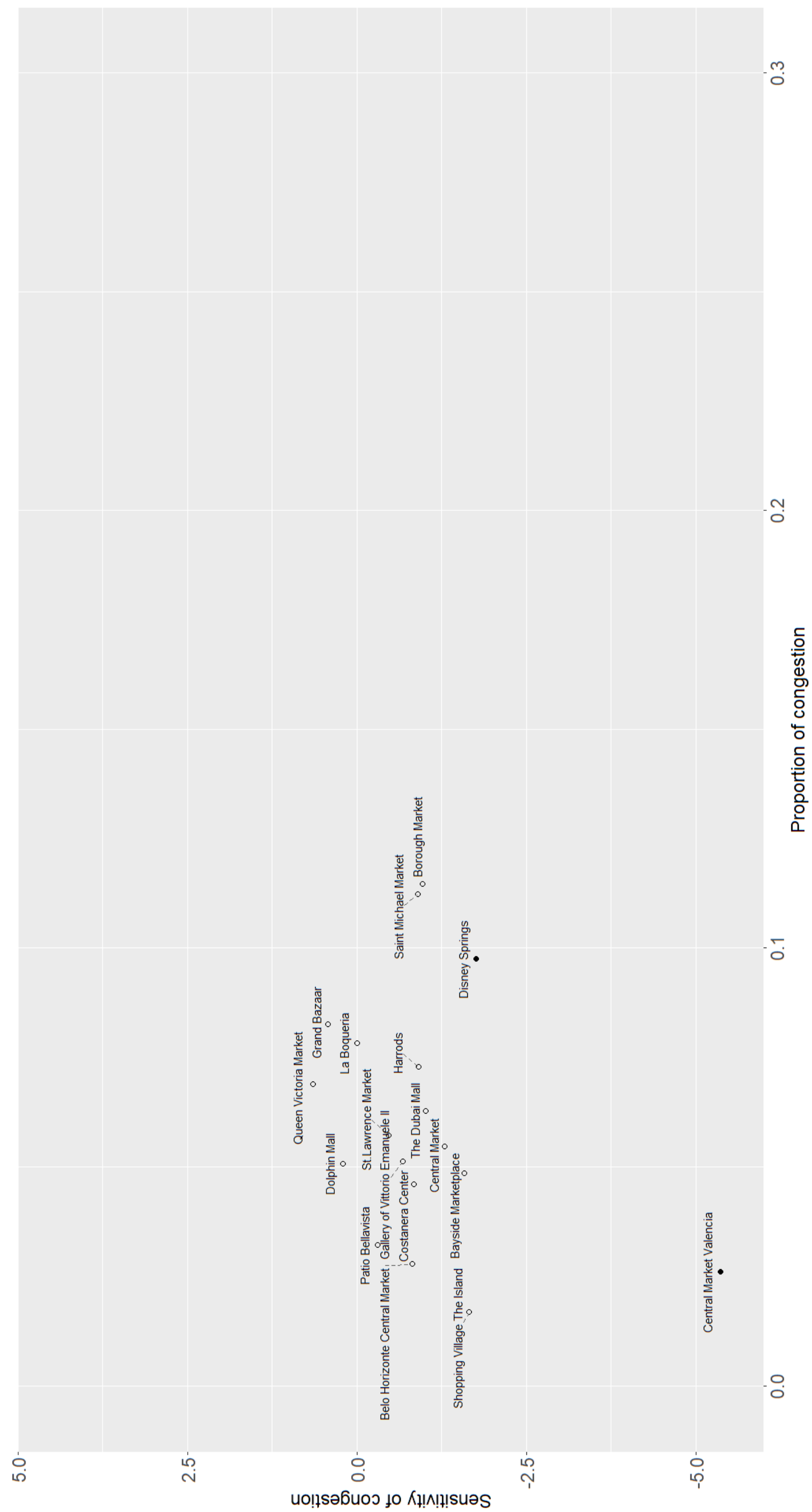

Figure 20 Top 500 Attractions. Congestion Chart for Shopping Areas (detail of Figure 5).

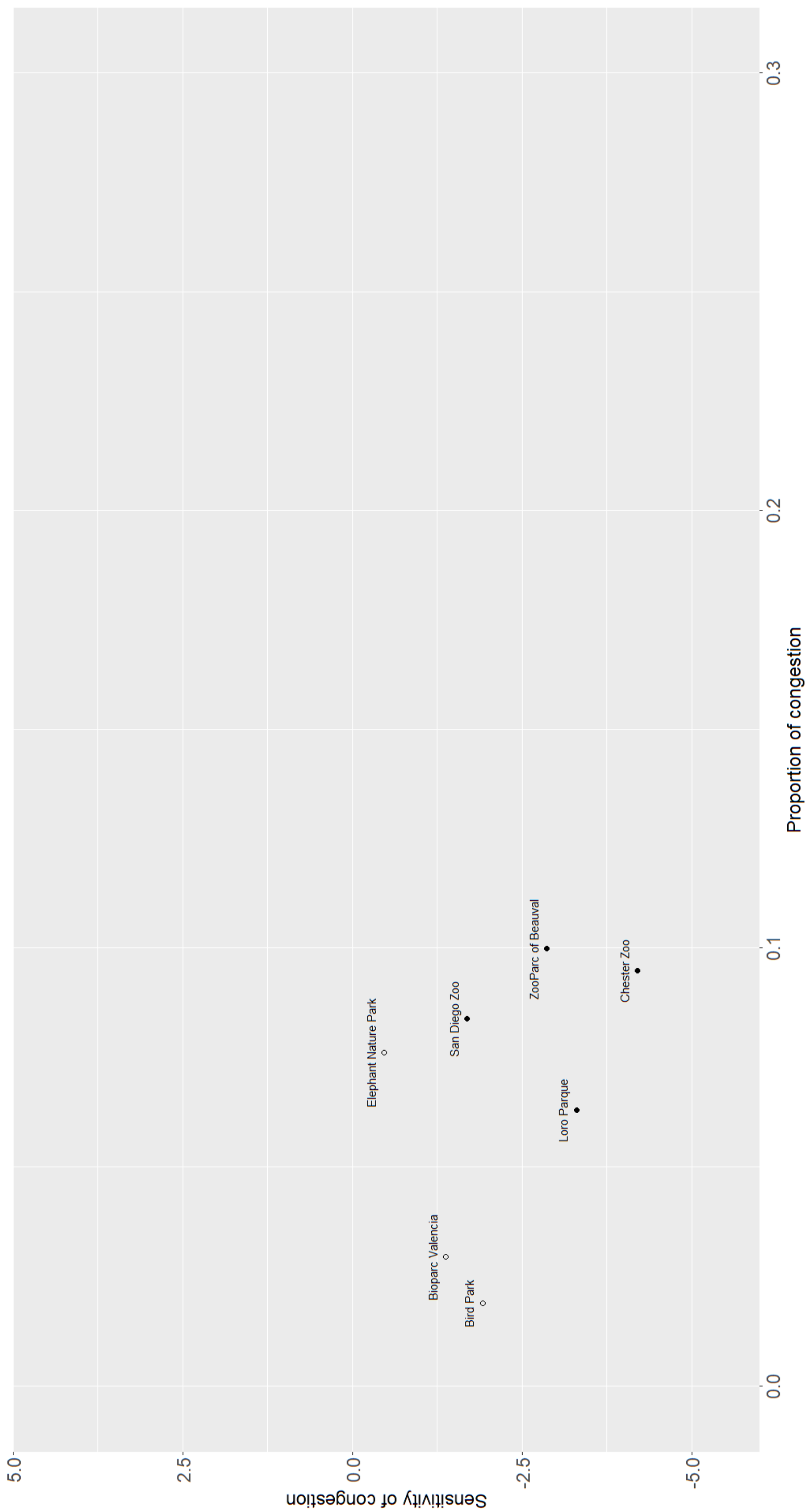

Figure 21 Top 500 Attractions. Congestion Chart for Zoos (detail of Figure 5).