You Can't Sit with Us: How Locals and Tourists Compete for Amenities in Paris*

Vladimir Avetian[†] Stefan Pauly[‡]

September 30, 2023

Abstract

Tourism in cities creates social interactions among people from distant cultures within limited space. How does the influx of tourists affect locals' satisfaction with amenities? Using data on restaurant reviews, we construct a panel of tourist presence in Paris. Based on two unanticipated drops in tourism – the November 2015 terrorist attack and the COVID-19 pandemic – we document that tourism reduces Parisians' satisfaction with restaurants. We find that social frictions, like xenophobia towards tourists, drive our results. As tourist numbers declined, explicit complaints about tourists in reviews decreased, while other complaints remained unaffected. Locals are least satisfied with dining among tourists from countries with weak social ties to France. Tourists are not affected by the presence of other tourists.

Keywords: Amenities, Discrimination, Tourism

JEL classification: O18, L83, J15, Z30

*The authors thank Eric Brousseau, Thomas Chaney, Pierre-Philippe Combes, Ruben Durante, Ilya Eryzhenskiy, Roberto Galbiati, Sergei Guriev, Andrei Kalk, Pamina Koenig, Benjamin Marx, Florian Oswald, and Ekaterina Zhuravskaya for their useful comments and suggestions. We are grateful to all the participants of the 16th North American Meeting of the Urban Economics Association, the 2022 SMU-Jinan Conference on Urban and Regional Economics, the PhD-Economics Virtual Seminar, and seminars at Sciences Po and University Paris-Dauphine - PSL. Special thanks to Lena Nikulina for suggesting the title.

[†]University Paris-Dauphine - PSL, *Chair of Governance and Regulation*: `vladimir.avetian@dauphine.psl.eu`

[‡]Sciences Po: `stefan.pauly@sciencespo.fr`

1 Introduction

"Are there too many tourists in Paris?" – was the title of a conference organized by the city hall of Paris on June 24, 2019. Experts and officials agreed that overtourism in Paris has not yet reached the same scale as in Amsterdam or Barcelona, but they admitted that "rapid and poorly regulated growth" of tourism can be harmful to the city.[1] There were reasons for concern. The number of foreign tourists to France had more than doubled over the previous 15 years. In 2019, France was the most visited country in the world, and Paris was the third most visited city. During that year, 35.4 million tourists stayed in the city's hotels, which is approximately 16 times more than the population of the city.

In the years preceding the pandemic, concerns about tourism became common in Europe.[2] Anti-tourist protests took place in Barcelona, San Sebastián, Mallorca, Venice, and other European cities. Anti-tourist graffiti, typically saying "tourist go home", were spreading across cities, including Paris. However, during the summer of 2020, there were no crowds of tourists in the major European cities. The problem of overtourism raised at the city hall conference faded into the background, when the COVID-19 pandemic and the stringency measures, imposed by governments, disrupted tourist inflows. This caused, as was coined by the World Tourism Organization, "the worst year in tourism history".

The unexpected shock in tourism created an opportunity to explore the question: "What would life be like for Parisians if there were no tourists?" During the summer of 2020, Parisians experienced a notable absence of tourists, while restaurants and other urban amenities remained accessible – they were kept open artificially through heavy government subsidies. This provides a unique setting to study demand-related factors without an endogenous adjustment of supply.

In this paper, we aim to estimate the effect of tourism on the subjective quality of life of residents and their satisfaction with urban amenities. Our approach is made possible by the availability of highly granular data on restaurant reviews, which we collected from

[1] The World Tourism Organization (UNWTO) defines overtourism as "the impact of tourism on a destination, or parts thereof, that excessively influences perceived quality of life of citizens and/or quality of visitor experiences in a negative way" (Carvão et al., 2018).

[2] See this Guardian article about anti-tourism protest in Europe: https://www.theguardian.com/travel/2017/aug/10/anti-tourism-marches-spread-across-europe-venice-barcelona (last retrieved September 19, 2023).

Tripadvisor – the platform that aggregates user-generated content on restaurant and other travel experiences. We view this data as a "digital footprint" of urban consumption. First, we use this data to construct a tourism measure at the restaurant level. Second, exploiting information from users' profiles, we identify reviews made by locals, and track individual reviewers over time. Third, we use ratings and texts of reviews left by Parisians as an indicator of their satisfaction with urban amenities.

With this detailed data on urban consumption at hand, we examine two instances of a sudden, unexpected drop in tourist arrivals. As mentioned above, the first episode we study is the COVID-19 pandemic and the associated drop in cross-border travel. This resulted in a dramatic decline in the number of international tourists. The second episode we study is the drop in tourist arrivals caused by the November 2015 terrorist attacks in Paris. These tragic events were widely covered by international news media and in the aftermath Paris was temporarily perceived as unsafe by many. As a result, Paris saw a sizable decline in tourists arriving from abroad, but the impact was smaller than that of COVID-19. The fact that some tourists were still present allows us to test whether tourists themselves are affected by other tourists.

We employ a difference-in-differences strategy, by classifying restaurants into treatment and control groups, based on their proportion of tourist customers prior to the exogenous decline in tourism. The outcome variable is the average rating by residents for a restaurant in a given month. We incorporate restaurant and month-neighborhood fixed effects and estimate the effect of the drop in tourism on residents' satisfaction with "tourist" relative to "non-tourist" restaurants. Our main specification thus addresses a series of potential endogeneity concerns. Both the COVID-19 pandemic and the November attacks could have directly affected residents' satisfaction and behavior. For instance, once the lockdown was lifted, residents might have felt a renewed appreciation for meeting in public spaces or may have been driven to support restaurants during these challenging times. By focusing on the differential between initially more and less touristic places within the same neighborhood over time, we can rule out these confounding factors.

We find consistent results for both natural experiments: a drop in tourism led to an increase in residents' satisfaction with urban amenities. For the shock brought about by the pandemic, the estimated effect is larger than that induced by the terrorist attack.

This aligns with the more substantial decline in tourist arrivals during the pandemic. The estimated effects are statistically significant, and the magnitude is meaningful. Focusing on the COVID-19 results, locals' satisfaction with the average restaurant in the most touristic neighborhoods increases by around 7%.

We repeat this exercise at the review level. This allows us to include user fixed effects. This accounts for changes in the composition of reviewers. In addition, we interact user fixed effects with a post-dummy, thereby comparing within-user-period differences in ratings of more and less tourist restaurants. If a third factor caused users to go to more tourist places and leave higher ratings, this is controlled for.

To complement and validate our amenity measure, we matched restaurant data with the number of complaints on the crowd-sourced platform DansMaRue provided by the city hall of Paris. This serves as another source of information on residents' quality of life. The platform allows users to report any problems related to public spaces, such as abandoned waste, graffiti, or illegal posters, using a mobile application or a website. Using this alternative measure of satisfaction with amenities, we find that the number of complaints around restaurants increases with tourism.

To further validate our findings, we conduct a placebo test by employing the same design and estimating the effect of the shock on ratings among non-Parisians. As expected, we do not observe a significant effect of the presence of tourists on tourists' satisfaction with urban amenities. It is worth noting that this test can only be performed for the shock associated with the November 2015 attacks, as there were not enough tourist customers in the aftermath of the first COVID-19 lockdown.

Our main results are robust to various modeling choices we made. We present similar findings obtained from specifications with different fixed effects, time periods, and alternative tourism measures.

In the second part of the paper, we consider three mechanisms that could explain our findings: overcrowding, supply-side changes, and social frictions such as xenophobia towards tourists.

Overcrowding, or congestion, happens when the number of people exceeds the capacity of the city's infrastructure. Tourist inflows are highly localized and concentrated: the geographic distribution of tourist arrivals is uneven both between and within cities.

Overcrowding can disrupt transportation, increase pollution, generate queues, and reduce the individual value of existing amenities.

Supply-side changes can be described by businesses, including restaurants, adapting based on the composition of their customers. For example, tourist-oriented neighborhoods have a higher proportion of French cuisine restaurants and fewer diverse cuisine options, reflecting horizontal differentiation (see Appendix Figures A.6 and A.7). Since this paper emphasizes short and medium-run effects, our main analysis centers on vertical differentiation – the way tourism can potentially influence restaurant quality. Tourists are typically one-time customers, whereas residents are repeat customers. With a majority of potential consumers being tourists, suppliers may have diminished incentives to maintain consistent quality.

The social frictions mechanism is related to locals' attitudes. Tourism adds another layer of diversity to cities, in addition to the racial and ethnic diversity already present among residents. Similar to racial segregation (Davis et al., 2019), the consumption of tourists is also segregated, with most consumption occurring in proximity to tourist attractions. Therefore, residents may have a (discriminatory) distaste for consuming together with tourists.

We perform several exercises to test the mechanism. First, we rely on text-as-data methods. The growing literature in economics uses these methods to enrich traditional data sources with new variables (Ash and Hansen, 2023). In our case, we want to analyze the texts of restaurant reviews to better understand the subjects of complaints. For this purpose, we identify five common motives present in the reviews: complaints about tourists, low food quality, excessive prices, long waiting times, and noise. We assign reviews to topics using two common approaches: dictionary-based methods and word embeddings.

The dictionary, or rule-based approach, has been widely used in the economic literature historically. It involves the use of pre-selected keywords and wildcards to label data. However, it is rarely possible to construct a complete and non-arbitrary dictionary. Word-embedding is a family of techniques used by computational linguists to represent words as vectors in a geometric space, where the position of the vectors reflects the semantics of the words. Semantic similarity between words can be measured using cosine distances. In this study, we use the *word2vec* approach developed in Mikolov et al. (2013), and we mostly follow the procedure of Gennaro and Ash (2022), who performed rhetorical analysis of congress speeches in the US. The drawback of this method is that semantic relationships may capture unrelated concepts or dimensions of meaning, and that *word2vec* performs poorly with negations.

Both methods yield similar results. We find that tourism increases complaints directly related to tourists, whereas it does not affect complaints related to poor quality of food, excessive prices, long wait times, and noise. If the presence of tourists in a restaurant does not impact food quality or other factors, yet reviews drop solely due to the presence of tourists, it could imply a general aversion among residents to dine alongside them. This is also in line with the fact that tourists themselves are not affected by the presence of other tourists.

Finally, relying on a proxy of social connectedness between countries derived from Facebook data, we find that restaurants with a clientele that has more distant connections to France see a larger increase in its rating post-lockdown. This suggests that Parisians are less bothered by tourists from countries with which they have strong social ties. Furthermore, the text analysis result is stronger for restaurants having consumers from socially distant countries.

We contribute to the literature on social frictions, discrimination, and segregation in cities by providing evidence of social frictions in the tourist-resident dimension. This complements more traditional results focusing on residents of different races. For example, Davis et al. (2019) show that urban consumption is segregated along racial lines in the US, Algan et al. (2016) study the effects of ethnic diversity of residents on social cohesion at the housing block level. In contrast, we focus on tourism, itself being an important source of social interactions in the city. Given that tourism can be viewed as a facet of international trade, we also contribute to the literature that shows that cultural, social, and ethnic proximity can impact trade patterns (Bandyopadhyay et al., 2008; Guiso et al., 2009; Macchiavello and Morjaria, 2015).

Despite the heated public debate surrounding "overtourism", there is limited causal evidence on the topic. We provide novel evidence that urban tourism has a negative effect on residents' satisfaction with amenities. Thus, we complement an existing literature on urban tourism. Tourism has been shown to have a substantial positive economic impact (Faber and Gaubert, 2019). Lanzara and Minerva (2019) demonstrated a positive correlation between an increasing number of tourists and the number of restaurants and bars. Hidalgo et al. (2022) show that the higher numbers of Airbnb arrivals cause an increase in the number of restaurants. On the other hand, Airbnb has been shown to increase rents (Garcia-López et al., 2020). Finally, Besley et al. (2020) demonstrate how terrorism can negatively impact tourism, supporting our decision to examine terrorism-induced shocks.

We also add to the nascent literature on tourism and amenities by underscoring the significance of social frictions. Almagro and Domínguez-Iino (2022) study how amenities and location sorting by residents endogenously adjust to a large increase in tourist demand, focusing on the city of Amsterdam. In contrast to this paper, we focus on how tourism affects satisfaction with existing amenities rather than how amenities adjust to tourism demand, thereby affecting locals. Allen et al. (2020) study the effects of tourism on residents' welfare in Barcelona using data on consumption patterns and find that tourism expenditure crowds out expenditure by locals. We use detailed restaurant-level data to show how sharing the same amenities with tourists can affect locals' reported quality of life and are able to distinguish different mechanisms through which this effect operates.

More generally, our paper builds on the literature emphasizing the importance of amenities. In their seminal paper, Glaeser et al. (2001) explore the role of cities as centres of consumption. They show that high-amenity cities have been growing faster than low-amenity cities, highlighting the importance of amenities for location choices. Generally, on the importance of urban amenities for attracting residents, see also Carlino and Saiz (2019), Lee (2010), and Couture and Handbury (2020). Finally, we contribute to the diverse literature on the COVID-19 pandemic and its interaction with the city (Althoff et al., 2022; Couture et al., 2022; Coven et al., 2023; De Fraja et al., 2020; Gupta et al., 2022; Miyauchi et al., 2021).

The remainder of the paper is organized as follows. In Section 2, we provide a background on the two tourism shocks that we study. Section 3 describes our main data sources. We explain our empirical strategy in Section 4 and present our main results in

Section 5. Section 6 provides robustness checks and additional results. In Section 7, we explore the potential mechanisms behind our main result. Section 8 concludes the paper.

2 Two Episodes of Sudden Tourism Decline

In this paper, we examine two exogenous shocks in tourism that occurred during distinct periods, both of which took place when *Tripadvisor* was in widespread use in Paris. Both shocks provide complementary insights into the impact of tourism on residents' satisfaction with amenities.

2.1 COVID-19 in Paris

The first restrictions related to COVID-19 were implemented in early 2020. On March 12, President Emmanuel Macron announced in a televised address that all schools and universities across France would close. The very next day, on March 13, 2020, Prime Minister Édouard Philippe declared the closure of all pubs, restaurants, cinemas, and nightclubs. After three months of strict lockdown measures, cafes, restaurants, and pubs reopened in Paris on June 14.

While the restaurant sector began its return to normalcy, tourism continued to suffer heavily due to the global pandemic. The Île-de-France region, which encompasses Paris and its suburbs, was particularly hard-hit. Compared to July 2019, overnight stays in its hotels in July 2020 decreased by 70.8%.[3] Subsequent months witnessed a comparable downturn in the hospitality sector. This decline was notably accentuated among tourists not residing in France. In 2020, France experienced a 71.8% reduction in non-resident overnight stays compared to 2019, whereas overnight stays by residents decreased by only 10.5%.

The drop in tourism and COVID-related restrictions led to a large shock to restaurant revenue. In response to this looming crisis, the government implemented safety nets in place that ensured the survival of most affected businesses. Bankruptcies actually declined dramatically during the pandemic. According to the Ministry of Finance, in

[3]See this article by the French national statistical agency INSEE: https://www.insee.fr/fr/statistiques/5369851 (last retrieved on September 15, 2023).

the food service sector, business closures were down 57% from March 2020 to October 2021 compared to the pre-pandemic period.[4] The composition of restaurants thus stayed remarkably stable over this period of changing demand. This provides us with an ideal setting to study the effects of competing for access to existing amenities.

To further illustrate the nature of the shock, Figure 1 shows the number of reviews written in French and other languages, separately. The beginning and the end of the "first-wave" lockdown imposed by the French government are marked with a blue dotted line. During the lockdown both French and non-French reviews dropped to near zero. Then, starting in June, French reviews revived, but foreign reviews remained on a negligible level. The observational period ends with both French and non-French review numbers going back to zero due to the introduction of a second wave of restrictions. This demonstrates that demand by tourists remained low during the summer after the first lockdown while locals quickly returned to restaurants.

2.2 November 2015 Paris attacks

On November 13, 2015, Paris was hit by a series of terrorist attacks. The gunmen fired at civilians in multiple locations across the city, leaving 130 dead. This gruesome attack immediately captured both national and international media attention. Detailed accounts of survivors were quickly spread across the globe. This indiscriminate attack on civilians made Paris an unsafe place to visit in the eyes of some and led to a substantial decline in tourism. For example, overnight stays in Paris by visitors from foreign countries dropped by 9.8% year-on-year in the fourth quarter of 2015.[5]

Figure 2 shows the decline in non-French reviews after the attacks, with no discernible impact for French reviews. The November attacks thus lead to a sudden decline in tourism. However, tourism did not decline as sharply as during the pandemic. This allows us to see how satisfaction with amenities evolved among those tourists that came despite the shootings.

[4]See this report for further details: https://www.tresor.economie.gouv.fr/Articles/2022/01/18/business-failures-in-france-during-the-covid-19-crisis (last retrieved on September 15, 2023).

[5]See this article published by the French national statistical agency INSEE: https://www.insee.fr/en/statistiques/2011367 (last retrieved September 19, 2023).

3 Data

In this section we first discuss our primary dataset on restaurant reviews collected from the website *Tripadvisor*, describe our measure of tourism and briefly introduce the text analysis we perform on the review texts. Finally, we describe additional datasets from other sources that we use.

3.1 Tripadvisor Data

Tripadvisor is a user-generated social media review site, which publishes user reviews on restaurants, hotels and other attractions. We collected data on all Parisian restaurants that were listed on the site on November 17, 2020.[6] We obtained information on restaurant characteristics, such as the type of cuisine and the address, and individual review data, including the review's date, text, language, user, user location and rating. We geocode restaurants' addresses. We leverage the data on review's language and user location to separate consumption of residents and tourists. As a result we construct a unique and highly detailed panel that reflects the city's restaurant consumption across space and time. French users began adapting the platform in 2007, and their usage peaked in 2017.

The entire set of Parisian reviews that we collected from 2008 to 2020 consisted of around 2 million reviews for approximately 15,000 restaurants, cafes, and bars. We then sampled and preprocessed them for further analysis.

3.2 Measuring Tourism

In this paper we use review data to construct a highly granular measure of tourism at the restaurant level. Importantly, it gives us an indicator of where tourists consume in the city rather than where they stay. Our preferred proxy of tourism is constructed as a share of reviews written in languages other than French. In Section C.1 in the Appendix we repeat our analysis using an alternative measure of tourism based on users' home locations.

The Figure 3 shows a map of our tourism measure. A lighter color indicates a higher

[6] In this analysis, we restrict ourselves to restaurants located in Paris *intra-muros* – the city of Paris that consists of 20 municipal arrondissements and excludes the surrounding Greater Paris area.

share of non-French reviews. As expected, restaurants with the highest levels of tourism are located in the areas known for Paris' major attractions: the Eiffel tower, Montmartre, Notre-Dame de Paris and the Arc de Triomphe.

To validate our proxy for tourism more formally, we use data from the *Enquêtes de fréquentation des sites culturels* (Attendance surveys of cultural sites) provided by the *Observatoire économique du tourisme parisien* (Observatory of the Parisian tourism economy). This survey details the proportion of all tourists coming to Paris who visit various tourist attractions. We consider tourists visiting from 2015 to 2019 and geocode the 18 attractions that are located intra-muros contained in the survey. Then, we construct a measure for demand by tourists that follows the market access framework widely used in the economic geography literature:

$$\text{Tourist Access}_i = \sum_j \frac{\text{Visitors}_j}{\text{Distance}_{ij}},$$

where we are implicitly assuming a distance elasticity of tourist consumption trips of -1. While we are not aware of a paper estimating this parameter specifically for demand by tourists, Miyauchi et al. (2021) look at the distance elasticity of location choice for consumption trips. They find a value of -1.09 and thus close to -1.

Next, we correlate our tourism proxy with the tourist demand measure. As Figure A.4 shows, we find a strong positive correlation between the two (the R^2 of a linear regression is 0.19). The correlation is robust to controlling for neighborhood fixed effects, meaning that, even after controlling for a relatively fine-grained spatial unit, the remaining variation in our tourism proxy is correlated with tourist access (see Table D.1). Together, this shows that our proxy for tourism correlates strongly with other, external measures of tourism.

Finally, to further corroborate our proxy for tourism, we rely on user location information. In particular, we compute the share of users by restaurant who indicate a location in a country other than France. As Figure A.5 shows, the two measures are highly correlated (the R^2 of a linear regression is around 0.77).

3.3 Content of Reviews

We perform text analysis of reviews to better understand users' concerns. We distinguish five potential disamenities we want to test for: discussion on the presence of tourists themselves, concerns about low food quality, high price, long waiting time and noisy environment.

To identify these concerns in the text, we employ two established approaches from the literature: a dictionary-based method and word embeddings. In the dictionary-based method, we pre-select a set of words and phrases associated with the relevant topics. For word embeddings, we adhere to the technique proposed by Gennaro and Ash (2022), which involves calculating cosine distances between the related dictionary centroids and the centroids of text reviews in word2vec embeddings. Each approach has its own advantages and drawbacks. In this section, we detail these methods and explain how we constructed each measure.

3.3.1 Dictionary-Based Approach

The mapping of review texts to topics is achieved through manually constructed dictionaries. The construction procedure unfolds as follows. First, we examined approximately one thousand randomly selected reviews to identify words that unambiguously pertain to the topic. Second, we validated these terms by searching for counter-examples within the corpus to highlight the "false positives" – reviews where these terms appear, but are not genuinely related to the topic. Third, we augmented our dictionary to include common misspellings and partial forms of the chosen terms. Lastly, we assembled a list of 'minus' phrases to ensure that expressions like *"pas cher"* (not expensive) are not mistakenly flagged as *"cher"* (expensive).

In essence, our methodology primarily minimizes *false positives* (the risk of erroneously attributing a text to a topic when it is not pertinent), but it does not actively reduce *false negatives* (the chance of overlooking a text's relevance to a topic). The concise version of our dictionary (excluding misspellings and variations) can be found in Appendix Table E.1. The summary statistics of the topics are presented in Appendix Table E.3. It is noteworthy that all topics manifest with relatively comparable frequencies (between 2% and 6%).

3.3.2 Word Embedding

To complement our dictionary-based measure, we adopt the word2vec method (Mikolov et al., 2013). This model represents words as vectors, often referred to as embeddings, that refine the extensive and sparse co-occurrence information throughout the corpus into more concise, lower-dimensional forms. In this generated space, the "semantic" distances between words can be assessed. This technique has been effective in many social science applications. Gennaro and Ash (2022) outline a method for using word2vec in analyzing economic data. We follow their guidelines for preprocessing, but make an exception for parts of speech selection, as we believe our topic might require the inclusion of additional linguistic elements beyond just verbs, subjects, or objects.

This approach offers the advantage of extracting more nuanced information from texts than just relying on the presence of predefined keywords. Considering that in many instances the initial set of keywords can be arbitrary, this method aids in formulating a more comprehensive measure. However, there are intrinsic limitations associated with word2vec. For instance, it struggles with polysemic words – words that have multiple meanings depending on the context. Additionally, word2vec often fails to distinguish between negation and positive use, making it challenging to derive precise sentiment. Moreover, interpreting the continuous distances based on embeddings is not always intuitive. It is crucial to highlight that the seeding dictionary we employed to construct our embedding measure differs from the one used in our dictionary-based measures. In the case of embeddings, we were restricted to using individual words instead of phrases. For a detailed look at this seeding vocabulary, please refer to the Appendix Table E.2.

It is important to note, as evident from Figure 7, that our defined topics, measured using cosine distance in embedding, are correlated in a meaningful way. They can logically be grouped by considered mechanisms. For example, the most correlated variables are those related to long waiting times and noise; both can be associated with the overcrowding mechanism.

3.4 Data from the Application "DansMaRue"

Most of our analysis is based on the TripAdvisor data. To externally validate that the presence of tourists affects locals' satisfaction with amenities, we draw on an additional dataset from the application *DansMaRue* created by the Municipality of Paris. With the help of this application, citizens can register and geolocalise 'anomalies' observed in public space in Paris.[7] Users upload the complaints directly from their smartphones, specifying the location, date and the subject. The aim of the application is to improve the quality of Parisian public space by giving access of user-generated data on 'anomalies' to municipal service. The application was launched in 2012. For our analysis we focus on complaints about commercial activity which is the category most related to restaurant activity.

The high resolution of the data allows us to only consider complaints that are possibly related to a particular restaurant. We assign complaints to a given restaurant within a 100m radius.

3.5 Social Connectedness Index

Below we want to test whether the origin of tourists has an impact on locals' perception of them. To proxy for social (or network) proximity between foreign countries and France we rely on the Social Connectedness Index (SCI) published by Facebook.[8] It is based on the number of Facebook friendships between users located in a pair of countries. More precisely, it is computed as

$$\text{Social Connectedness}_{ij} = \frac{\text{FB Friends}_{ij}}{\text{FB Users}_i \times \text{FB Users}_j},$$

where FB Friends_{ij} are the number of friendships between users residing in countries i and j and FB Users_i the number of users in country i. The central premise of the index is that the 'network' connectedness between regions and countries acts as a measure of familiarity among their residents and the closeness of their social relationships. Social connectedness is influenced by cultural, geographical, political, and historical ties. For further details

[7] The set of potential 'anomalies' includes overflowing litter bins, illegal graffiti, abandoned objects, road damage and many others.

[8] The version we use dates from October 2021.

on the methodology see Bailey et al. (2018). Relying again on the information on users' origin, we compute the average social connectedness between the French population and the non-French customers of a particular restaurant.

4 Empirical Strategy

We employ a standard difference-in-differences framework at two different levels of aggregation to study the impact of the absence of tourists on locals' valuation of amenities. First, a restaurant-level approach gives us a broad picture of whether more and less touristic venues evolved differently over time. Second, review-level regressions allow us to asses whether the same users evaluated initially more touristic restaurant differently when borders were closed.

4.1 Restaurant-Level Analysis

At the restaurant level, we use the following specification:

$$\mathrm{Y}_{jt} = \beta \times \mathrm{Post}_t \times \mathrm{Tourism}_j + \gamma_j + \delta_t + \epsilon_{jt}, \tag{1}$$

where Y_{jt} is an outcome of restaurant j in month t. Our explanatory variable here and in the following specifications is $\mathrm{Post}_t \times \mathrm{Tourism}_j$. Post_t is a binary variable indicating whether month t belongs to the post-lockdown (or post-attack) period. $\mathrm{Tourism}_j$ indicates the extent to which restaurant j is frequented by tourists. It is measured by the share of non-Parisian reviews prior to a shock. We include restaurant fixed effects (γ_j) and month fixed effects (δ_t). In a more stringent variation of this specification we also include neighborhood-time fixed effects. This controls for any unobserved time-varying factors at the neighborhood level, such as an increased share of remote work during COVID-19, that may affect residential neighborhoods differently than the business district. Standard errors are clustered at the neighborhood level.

Below we will focus on one main outcome. We look at the average rating that restaurant j receives in month t, only looking at reviews by residents. Our hypothesis is that tourism lowers the utility residents derive from amenities. We thus expect $\beta > 0$.

4.2 Review-Level Analysis

At the review level, we use the following specification:

$$\mathrm{Y}_{ijt} = \beta \times \mathrm{Post}_t \times \mathrm{Tourism}_j + \gamma_j + \delta_t + \mu_i + \epsilon_{ijt}, \tag{2}$$

where Y_{ijt} is a rating by user i for restaurant j in month t. Our explanatory variable is the same as above. In addition to restaurant and month fixed effects (γ_j, δ_t), we also include user fixed effects, relying on within-user changes pre- to post-lockdown. Again, we cluster standard errors at the neighborhood level.

While including user fixed effects is already restrictive, identification can still come from comparing the magnitude of within-user changes across users, depending on whether they visited a touristic restaurant or not. If e.g. an increased life satisfaction post-lockdown and the propensity to visit more touristic restaurants were both determined by an unobserved third factor, our findings would be spurious. We thus, in a final step, interact user fixed effects with a post-lockdown dummy. This restricts identification to users who review at least two restaurants either before or after the lockdown. Intuitively, this specification captures whether the penalty for more tourist places decreased after the lockdown relying only on different ratings for more or less tourist restaurants by a user in the same period.

Our parameter of interest is β. Our hypothesis is that tourism is detrimental for residents' utility derived from a restaurant visit. Hence, we should observe that post-lockdown, when restaurants were open, but tourists were not present, initially tourist places start receiving higher ratings ($\beta > 0$).

5 Main Results

5.1 Pandemic-Induced Tourism Decline

Columns 1 and 2 of Table 1 show the results of estimating Equation 1 using the average monthly rating by Parisians at the restaurant level as the outcome variable and comparing the periods before and after the first COVID-19 lockdown. The sample period goes from January 2018 to November 2020. The results show that initially more touristic venues receive higher ratings by locals when tourists are no longer around. Importantly, as shown in Column 2, this effect is not driven by neighborhood-level trends as including neighborhood-time fixed effects only marginally changes the coefficient.

To better understand the magnitude of our estimates, consider a restaurant located in the area surrounding the Notre-Dame cathedral, a major tourist attraction. The average share of non-French reviews in this area is 64%. The estimates in Column 2 imply that locals' rating for this restaurant would increase in the absence of tourists by around 0.05 on a scale from 0 to 1, or by 7% relative to the mean. Tourism thus causes a substantial decrease in locals' satisfaction with amenities.

Table 2 shows the results of a user-level estimation (see Equation 2). Columns 1 to 4 in the first panel confirm our results at the user level, i.e. Parisians rate their experience higher in places previously frequented by many reviewers not from Paris. The coefficient is of similar magnitude as at the restaurant level. Importantly, in Columns 2 to 4 we include user fixed effects, exploiting within-user changes in behavior while holding fixed time-invariant individual characteristics, such as preferences for certain types of neighborhoods or restaurant types.

The nature of restaurant reviews does not allow us to control for user-restaurant fixed effects, since the vast majority of users rates a restaurant only once. However, we can allow for the user fixed effect to vary between the pre- and post-lockdown period. If, for example, there is an unobserved factor that causes both a mood shift among users and a systematic change in visiting more (or less) touristic restaurants, this could bias our results. Including a user fixed effect interacted with the post-lockdown dummy accounts for this possibility.

5.2 November 2015 Attacks

As argued above, the pandemic provides us with a sudden decline in tourism demand, while leaving existing amenities mostly intact. Still the pandemic may have affected restaurants in ways that are unobservable to us and correlated with our measures of tourism. For example, restaurants with larger outdoor facilities may have benefited most after the lockdown was lifted, as people continued to be cautious because of the risk to get infected. If the availability of outdoor facilities is correlated with our measure of tourism, we are wrongly attributing the observed changes in ratings and demand to tourism.

To alleviate concerns related to the specific nature of the pandemic, we instead use the terrorist attacks that took place on November 13, 2015 as an exogenous shock to tourism. Columns 3 and 4 in Table 1 display the results of estimating Equation 1 using reviews from November 2014 to November 2016.[9] We find that initially more touristic restaurants received better ratings by Parisians after the November attacks. Compared to the lockdown-related shock, the coefficient is smaller, which is in line with a lower drop in tourism arrivals than during the summer of 2020.

While we conduct restaurant level analysis with the terrorism shock, applying a review-level analysis here is too demanding. We provide the review level specification for the terrorist attack shock in Appendix Table B.2. The results are noisier. While we find a statistically significant effect in the first column, we estimate a positive, but insignificant coefficient when controlling for user fixed effects. As explained previously, the size of the tourism decline caused by the November was smaller, and likely dissipated faster as media attention faded.

Overall, this very different natural experiment lends support to our hypothesis that tourism negatively affects the quality of amenities as perceived by locals. The result does not seem to be driven by factors specific to the pandemic.

Finally, the November attacks allow us to look at the reaction of reviewers that are not from Paris. As shown in Columns 5 and 6 of Table 1, there is no effect on their ratings of touristic places. Externalities caused by tourism specifically affect locals. This suggests that general disamenities such as congestion are unlikely to be at play, but rather the presence of tourists themselves bothers locals.

6 Robustness & Further Results

In this section we first present results using the data on neighborhood complaints as a different measure of disamenities. Then, we show that our result is not specific to the pandemic-induced shock to tourism, not driven by pre-trends, not affected by spillovers and present minor robustness exercises such as different levels of clustering.

[9] We define tourism intensity based on data from 2014. November 2015 is dropped from the analysis and December 2015 onwards is defined as post-attacks.

6.1 Neighborhood Complaints

So far we have focused only on data coming from *Tripadvisor*. To provide further evidence that the lower influx of tourists improved locals' perceived satisfaction with local amenities, we analyze data on complaints registered within 100m of the restaurants in our sample by local residents (see Section 3 for a detailed description).

We estimate Equation 1, replacing the average rating of the restaurant with the number of complaints in the vicinity of a restaurant within a given month. As this is a count variable containing zeros, we use a Poisson model to estimate this equation.[10]

Table B.3 presents the results. We find that complaints around touristic restaurants decline relative to less touristic ones. Using the most conservative estimate in Column 2, complaints around a restaurant with an average share of tourists among its customers decrease by around 8%.[11]

The positive impact of a decrease in the arrival of tourists is thus not only reflected in restaurant ratings, but also confirmed by an entirely external data source, namely crowd-sourced complaints that are used to improve municipal services.

6.2 Testing for Pre-Trends

In order to asses the timing of the effect that we find, we estimate Equation 1 allowing for β to be time-varying. In particular, we estimate one coefficient per quarter and set the first quarter of 2020 as reference group. If the effect is driven by the sudden and unexpected absence of tourists due to the pandemic, we should observe no differential trends for more touristic restaurants prior to the outbreak of COVID-19. Figure 4 shows that this is the case. Prior to the COVID-19 outbreak coefficients are close to and not statistically different from zero. Then, in Q3 and Q4 of 2020 coefficients are positive and statistically different from zero. This lends further support to the interpretation that COVID-19 led to a shift in locals' ratings of initially touristic venues. Similarly, Figure 6 shows the the absence of pretrends at the review level and Figure 5 for the November 2015 attacks.

[10]More precisely, we use a Poisson Pseudo Maximum Likelihood estimator.

[11]We use the average tourism share of 31.6% and multiply it with the coefficient in Column 2.

6.3 Further Robustness Checks

The analysis is focused on tourists visiting a particular restaurant. We thus far have not tested if this effect spills over to restaurants located close by. In this case the effect of tourism would be further amplified. We thus include in our baseline specification, Equation 1, measures of how many tourists visit restaurants in the surrounding area. As Table 6 shows, using different distances, we do not find strong evidence for that. The impact of a reduced influx of tourists seems to be mostly limited to the restaurant itself.

In order to lend further credibility to our main result we perform several robustness exercises. First, we report our main result clustering standard errors at different levels. As Table C.5 shows, clustering at the neighbourhood level as done throughout our analysis is on the conservative side. Second, we use different measures of tourism. In Table C.4 we vary the period over which we compute the initial tourism share. Again, our results are robust to these different permutations. Third, we use the share of reviews left by non-Parisians instead of the share of reviews not written in French. As Table C.1 and Table C.2 illustrate, using this different proxy results in a qualitatively similar effect, both at the restaurant and at the review level.[12]

7 Mechanisms

To get at the mechanism, we look at the content of reviews. We rely on the text-based classification of reviews described in Section 3. This allows us to analyse, if reviews referring to certain topics are becoming more or less frequent. In particular, we estimate the following equation:

$$\text{Complaint on subject}_{ijt} = \beta \times \text{Post}_t \times \text{Tourism}_j + \gamma_j + \delta_t + \mu_i + \epsilon_{ijt}, \qquad (3)$$

where *Complaint on subject*$_{jt}$ indicates whether review left by user i for restaurant j in month t referring to a particular subject of complaint, such as overcrowding. The dependent variable is measured in two different ways: it is a dummy for the dictionary-

[12]Note that this measure likely also captures domestic tourism. Since travel restrictions mainly applied to international visitors, we focus on the share of non-French reviews. In addition, we do not observe the location of all reviewers.

based approach and a continuous variable between 0 and 1 for the word embedding approach. The rest of the specification is as described in the section related to Equation 2. We also estimate a restaurant-level version of this specification.

We start by looking at supply-side changes. As tourism declines, restaurants may change what they offer. Since locals are more likely to be repeat customers, businesses may invest more in quality. Similarly, as demand declines, prices may decrease, increasing locals' satisfaction with amenities. As Columns 2 and 3 in Table 3 shows, we find no evidence for less complaints about prices or poor quality.

Next, we look at another common disamenity associated with tourism, overcrowding. A decrease of people in the city due to the decline in tourism may simply lead to less congestion. To get at this we look at comments mentioning either a long waiting time or a noisy environment. As Columns 4 and 5 in Table 3 show, we find no evidence pointing in this direction. More touristic restaurants did not receive relatively less reviews mentioning a long wait or noise after the lockdown. We interpret this as congestion not being a major driver of our results.

Overall, we thus find no evidence for disamenities typically associated with tourism. This is in line with the fact that we find no effect of the November 2015 attacks on ratings of tourists. Congestion, high prices, and, maybe to a lesser extent, changes in quality are likely to be negatively perceived by tourists, too. The absence of an effect on ratings by tourists makes these mechanisms unlikely to be at play. Instead, this suggests that the presence of tourists affects locals differently.

Since we find no evidence for more classical disamenities, we next look for social frictions, i.e. a direct, taste-based aversion of locals against tourists. As the first column in Table 3 shows, the only reviews that explicitly mention tourists appear significantly less after the lockdown in initially touristic places. This suggests that it is something about the presence of tourists themselves rather than perceived overcrowding or decreases in quality.[13]

To further test whether social frictions are at play, we exploit the composition of tourists, which varies by restaurant. We test whether the increase in ratings is higher when the tourists are socially more distant to the local population. In particular, we exploit the

[13]Table C.3 replicates this result using the location-based tourism measure.

information on users' origin provided in their profile. This allows us to compute for each restaurant the share of reviewers from a given country of origin. We combine this with the Social Connectedness Index (SCI) to compute the average SCI between restaurants' foreign reviewers and France.[14]

If Parisians have a distaste for foreigners from less familiar countries, we should see a higher increase in satisfaction for restaurants with many visitors from these countries. We thus estimate the treatment effect separately for restaurants with above and below-median SCI value. Table 4 shows that the increase in ratings of touristic places is indeed driven by low-SCI restaurants. For example, in Column 3, when including month-neighborhood fixed effects, the treatment effect for high-SCI is close to and not statistically different from zero. The coefficient for low-SCI places on the other hand suggests that touristic, low-SCI restaurants increased their rating.

This evidence is thus consistent with social frictions. Locals are less bothered by tourists who are similar to them.[15] We can also see in Table 5 that, in a similar way, restaurants frequented by tourists from countries with a lower SCI index show a higher likelihood of complaints related to the presence of tourism. However, we do not observe any statistically significant results for other subjects of complaints.

8 Conclusion

This paper studies the impact of tourism on locals' satisfaction with a key urban amenity, restaurants. We construct a granular dataset of urban consumption, which allows us to introduce a restaurant-level measure of tourism. Exploiting two different exogenous declines in tourism, we find robust evidence that the absence of tourists increases locals' satisfaction with restaurants.

Using recently developed text analysis methods to analyse the content of reviews, we find no evidence that the effect operates through a reduction in overcrowding or through changes in prices or quality. Instead, the results point towards social frictions. Reviews

[14]See Section 3.5 for a description of the SCI.

[15]One concern might be that social connectedness is correlated with actual tourist arrivals from a country during the post-lockdown summer. However, the nature of the shock is such that arrivals from all countries drop to almost zero. Identification is thus almost entirely based on the pre-COVID exposure to tourism. In unreported results we control for differential changes in demand by nationality using a Bartik-style shock and find almost no change in our estimates.

that explicitly relate to the presence of tourists become less frequent. In addition, the negative effect of tourism is driven by restaurants where the initial tourist clientele was from countries that have few social ties with the French population. Finally, tourists themselves are not affected by having less tourists around them, lending further support to the existence of social frictions between tourists and locals.

While there is evidence that tourism can have adverse economic effects on locals, e.g. by increasing rents, our results suggest that at least some of the widely spread attempts to limit tourism may be driven by hostility towards foreigners. The physical presence of tourists creates social frictions and decreases locals' perceived quality of life. This calls for a careful evaluation of seemingly rational arguments that are used to curb tourism. We leave it to future research to study determinants of the degree of social frictions and how policy makers can try to reduce them.

References

Algan, Y., C. Hémet, and D. D. Laitin (2016). The social effects of ethnic diversity at the local level: A natural experiment with exogenous residential allocation. Journal of Political Economy 124(3), 696–733.

Allen, T., S. Fuchs, S. Ganapati, A. Graziano, R. Madera, and J. Montoriol-Garriga (2020). Is tourism good for locals? Evidence from Barcelona.

Almagro, M. and T. Domínguez-Iino (2022). Location sorting and endogenous amenities: Evidence from Amsterdam. Available at SSRN 4279562.

Althoff, L., F. Eckert, S. Ganapati, and C. Walsh (2022). The geography of remote work. Regional Science and Urban Economics 93, 103770.

Ash, E. and S. Hansen (2023). Text algorithms in economics. Annual Review of Economics 15.

Bailey, M., R. Cao, T. Kuchler, J. Stroebel, and A. Wong (2018). Social connectedness: Measurement, determinants, and effects. Journal of Economic Perspectives 32(3), 259–80.

Bandyopadhyay, S., C. C. Coughlin, and H. J. Wall (2008). Ethnic networks and us exports. Review of International Economics 16(1), 199–213.

Besley, T. J., T. Fetzer, and H. F. Mueller (2020). Terror and tourism: The economic consequences of media coverage.

Carlino, G. A. and A. Saiz (2019). Beautiful city: Leisure amenities and urban growth. Journal of Regional Science 59(3), 369–408.

Carvão, S., K. Koens, and A. Postma (2018). Presentation of UNWTO report "Overtourism: Understanding and managing urban tourism growth beyond perceptions".

Couture, V., J. I. Dingel, A. Green, J. Handbury, and K. R. Williams (2022). Jue insight: Measuring movement and social contact with smartphone data: a real-time application to COVID-19. Journal of Urban Economics 127, 103328.

Couture, V. and J. Handbury (2020). Urban revival in america. Journal of Urban Economics 119, 103267.

Coven, J., A. Gupta, and I. Yao (2023). JUE insight: Urban flight seeded the COVID-19 pandemic across the United States. Journal of Urban Economics 133, 103489.

Davis, D. R., J. I. Dingel, J. Monras, and E. Morales (2019). How segregated is urban consumption? Journal of Political Economy 127(4), 1684–1738.

De Fraja, G., J. Matheson, and J. Rockey (2020). Zoomshock: The geography and local labour market consequences of working from home. Available at SSRN 3752977.

Faber, B. and C. Gaubert (2019). Tourism and economic development: Evidence from Mexico's coastline. American Economic Review 109(6), 2245–93.

Garcia-López, M.-À., J. Jofre-Monseny, R. Martínez-Mazza, and M. Segú (2020). Do short-term rental platforms affect housing markets? evidence from airbnb in barcelona. Journal of Urban Economics 119, 103278.

Gennaro, G. and E. Ash (2022). Emotion and reason in political language. The Economic Journal 132(643), 1037–1059.

Glaeser, E. L., J. Kolko, and A. Saiz (2001). Consumer city. Journal of Economic Geography 1(1), 27–50.

Guiso, L., P. Sapienza, and L. Zingales (2009). Cultural biases in economic exchange? The Quarterly Journal of Economics 124(3), 1095–1131.

Gupta, A., S. Van Nieuwerburgh, and C. Kontokosta (2022). Take the Q train: Value capture of public infrastructure projects. Journal of Urban Economics 129, 103422.

Hidalgo, A., M. Riccaboni, and F. J. Velázquez (2022). The effect of short-term rentals on local consumption amenities: Evidence from Madrid. Available at SSRN 4000918.

Lanzara, G. and G. A. Minerva (2019). Tourism, amenities, and welfare in an urban setting. Journal of Regional Science 59(3), 452–479.

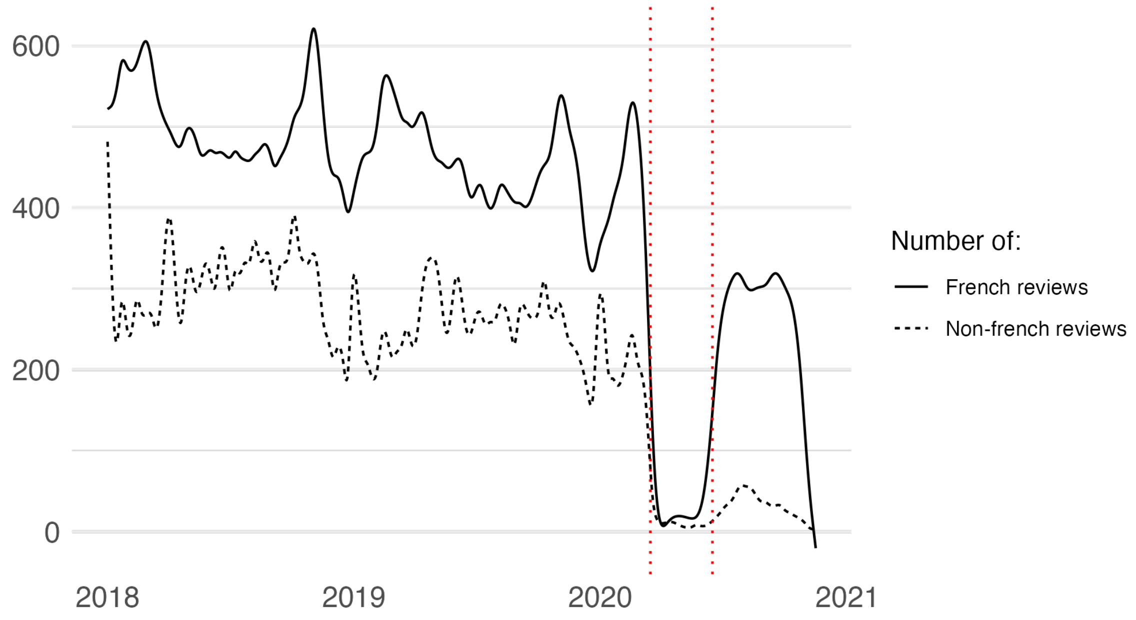
Lee, S. (2010). Ability sorting and consumer city. Journal of Urban Economics 68(1), 20–33.

Macchiavello, R. and A. Morjaria (2015). The value of relationships: evidence from a supply shock to kenyan rose exports. American Economic Review 105(9), 2911–2945.

Mikolov, T., I. Sutskever, K. Chen, G. S. Corrado, and J. Dean (2013). Distributed representations of words and phrases and their compositionality. Advances in Neural Information Processing Systems 26.

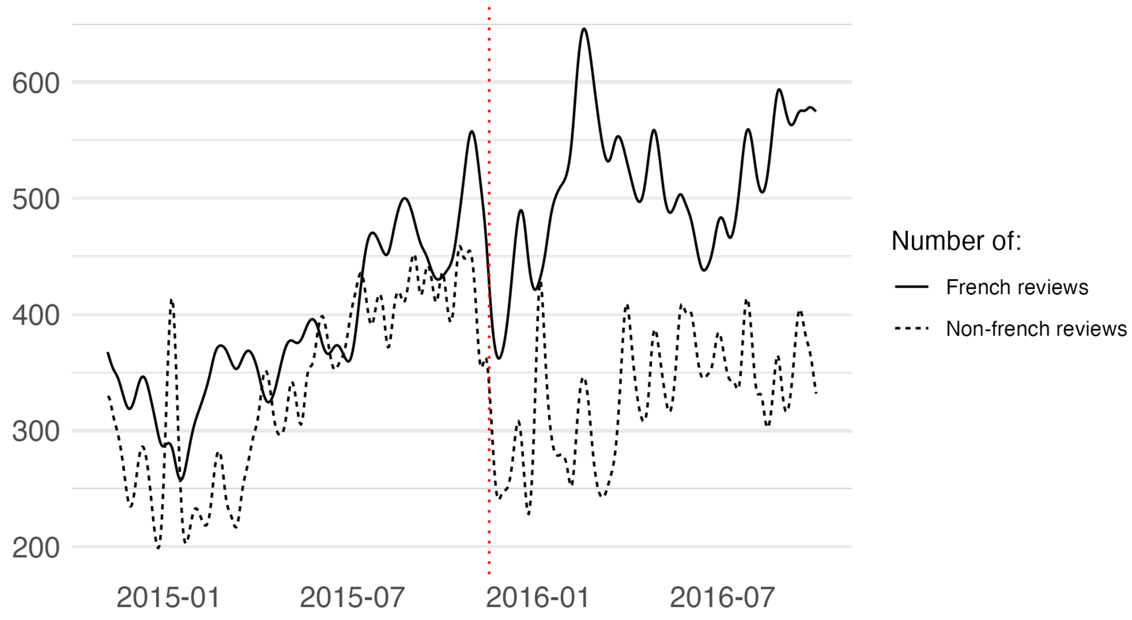
Miyauchi, Y., K. Nakajima, and S. J. Redding (2021). Consumption access and agglomeration: evidence from smartphone data. Technical report, National Bureau of Economic Research.

Figure 1: No Tourists in Summer 2020: French Restaurant Reviews Partially Recovered after the First Lockdown, while Reviews in Other Languages Remained Close to Zero



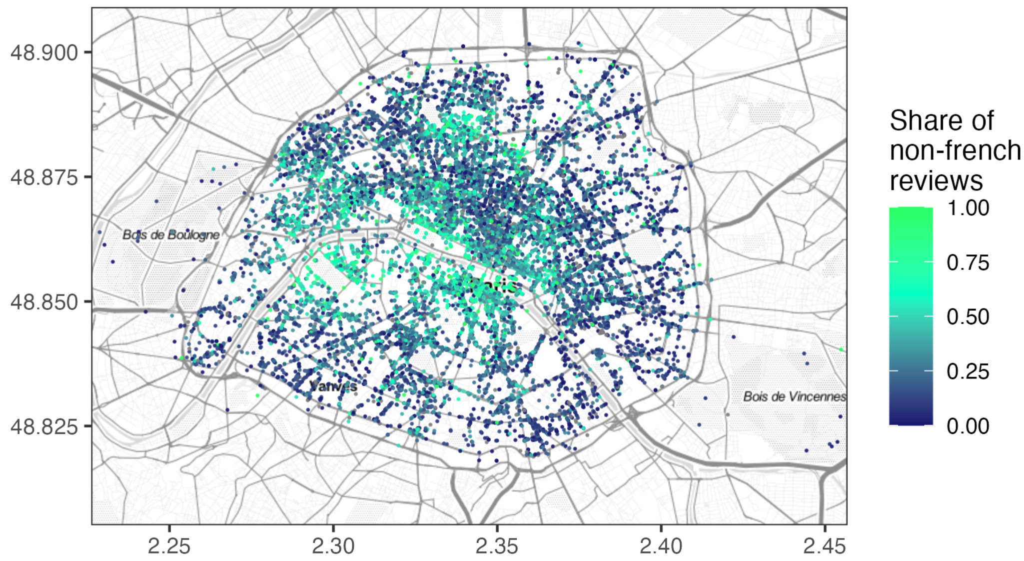
Notes: This figure illustrates the impact of the pandemic on restaurant reviews. Both lines correspond to the trends in the daily number of reviews, which are represented with smoothing splines, from January 2018 to November 2020. The solid line shows the French language reviews, while the dotted line represents non-French reviews. The two vertical red dotted lines mark the period of the first pandemic lockdown in Paris, which occurred between March and June 2020. During this period, restaurants were closed for visits. In the summer of 2020, the number of French reviews partially recovered, while non-French reviews remained close to zero.

Figure 2: After the November 2015 Terrorist Attack in Paris, the Number of Non-French Reviews Dropped



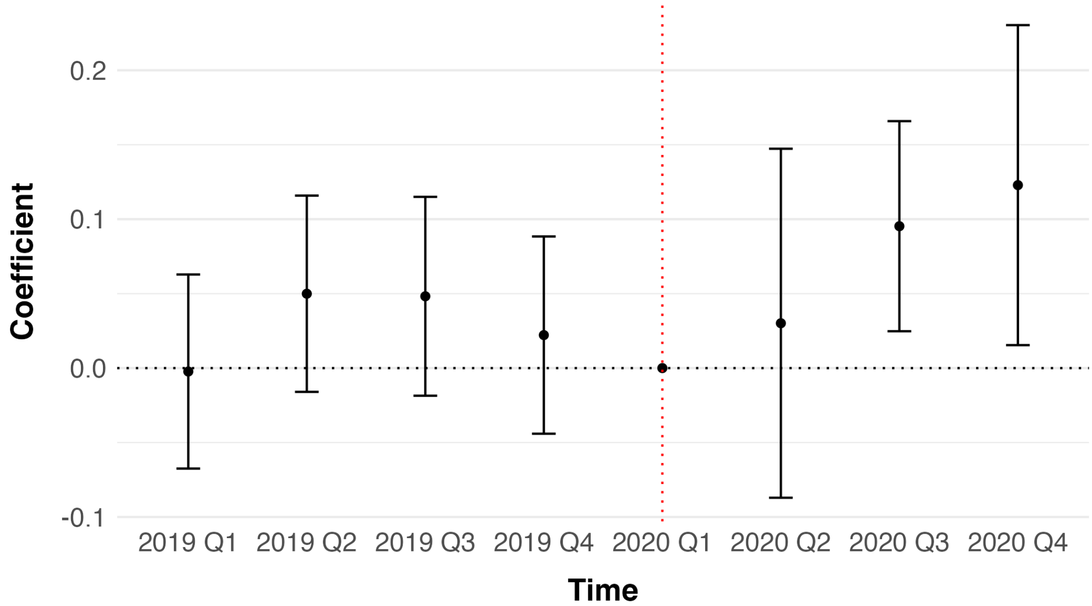
Notes: This figure illustrates the impact of the 2015 terrorist attacks on restaurant reviews. Both lines correspond to the trend in the daily number of reviews, which are represented with smoothing splines, from November 2014 to November 2016. The solid line shows the French language reviews, while the dotted line represents non-French reviews. The vertical red dotted line marks the day of the terrorist attack. After the attack, the number of French and non-French reviews diverged.

Figure 3: The Map of Restaurants, Categorized by the Share of Non-French Reviews, Corresponds to the Distribution of Major Tourist Attractions



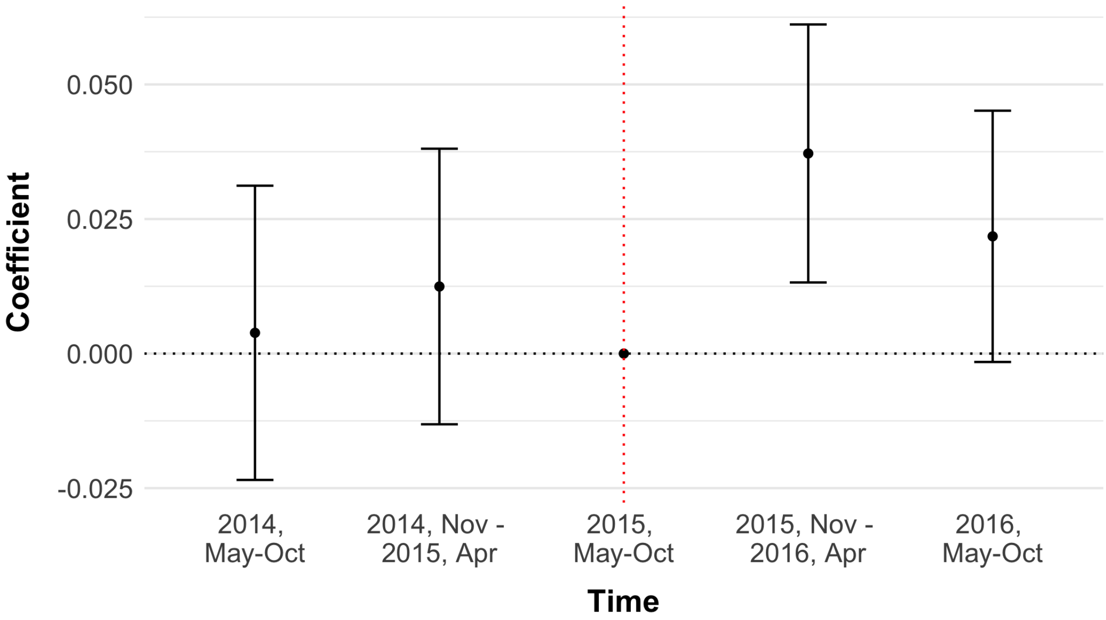
Notes: This map illustrates the distribution of restaurants based on the share of non-French reviews. Each point corresponds to a restaurant, and the color indicates the proportion of non-French reviews out of the total reviews prior to 2020. The color distribution indicates the level of tourist presence in the neighborhood, as the majority of restaurants with a high share of non-French reviews are clustered around major tourist attractions. The grid cell map, which mirrors the distribution of touristic restaurants, is depicted in Appendix Figure A.2.

Figure 4: Event Study Plot: Touristic Restaurants Have Relative Improvement in Ratings After Pandemic, Restaurant-Level Specification



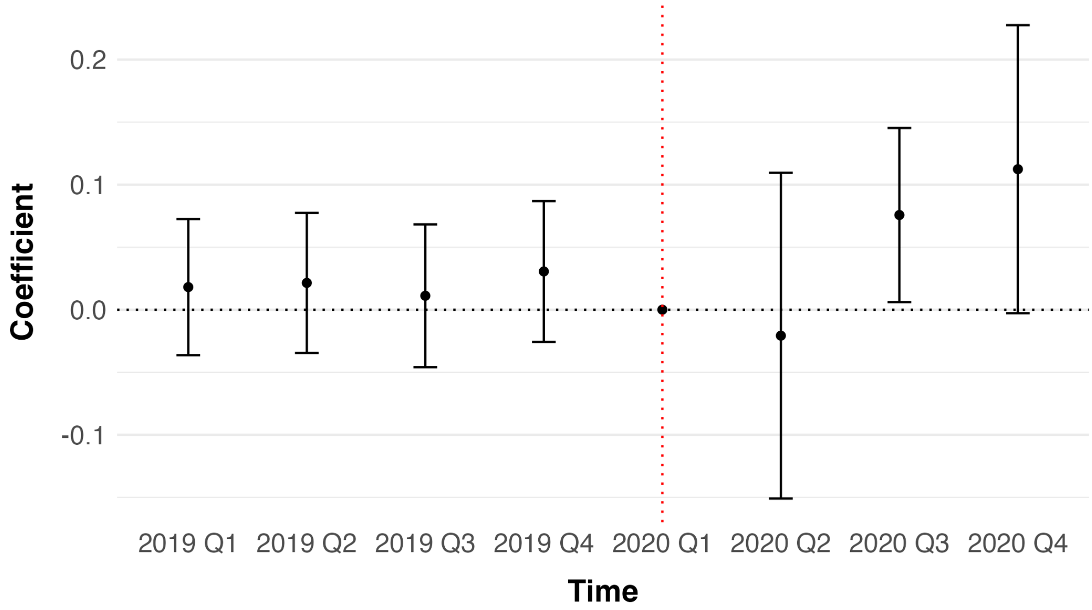
Notes: This figure displays the results of the event study concerning the pandemic-induced shock, where the unit of analysis is month $\times$ restaurant. Point estimates and 95% confidence intervals are provided. The first quarter of 2020 serves as the excluded time period. Standard errors are clustered at the quarter level.

Figure 5: Event Study Plot: Touristic Restaurants Have Relative Improvement in Ratings After November 2015 Attack, Restaurant-Level Specification



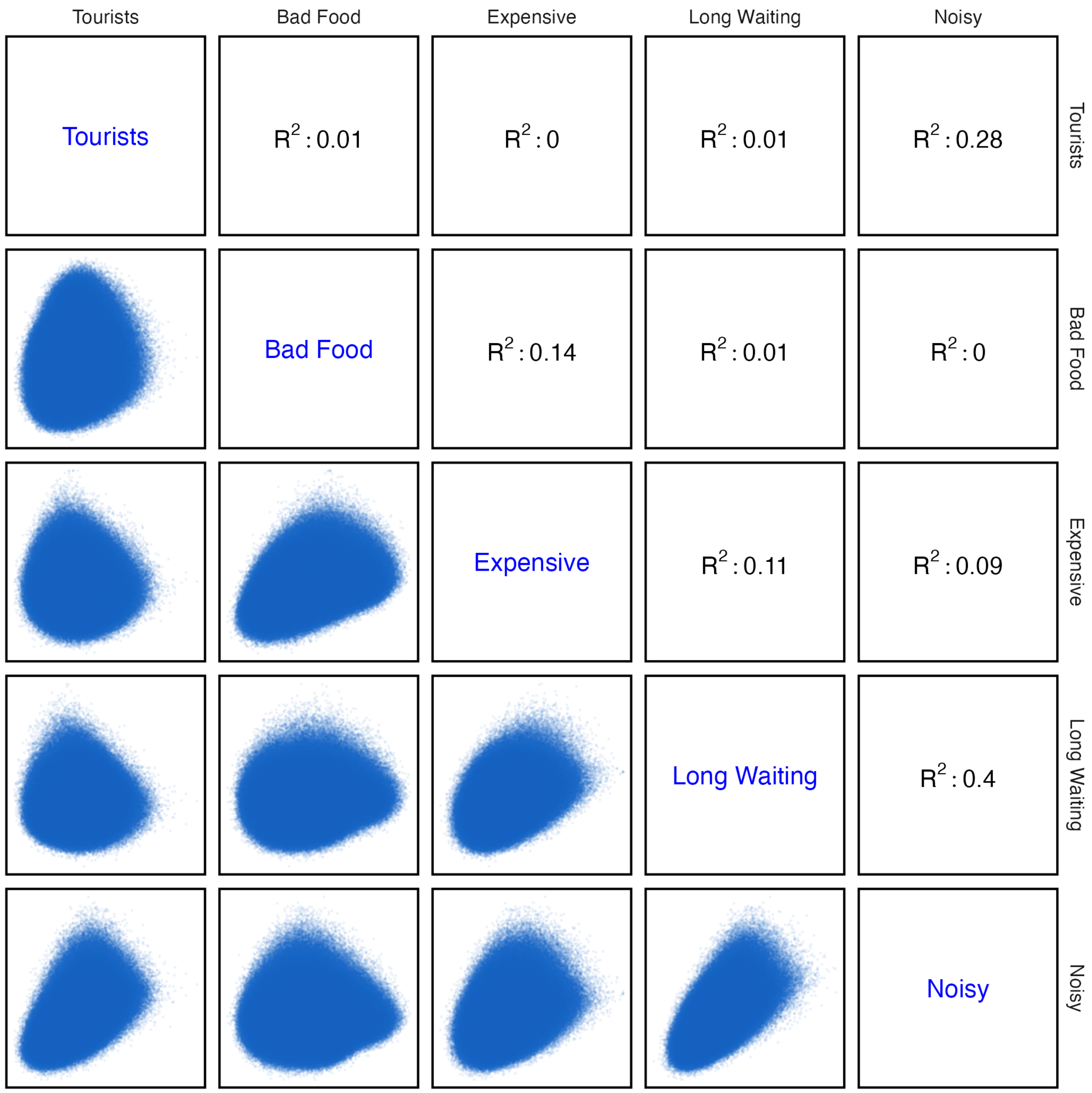
Notes: This figure displays the results of the event study concerning the terrorism-induced shock, where the unit of analysis is month $\times$ restaurant. Point estimates and 95% confidence intervals are provided. Points are grouped by half-a-year periods. The period from May to October 2015 serves as the excluded time period. The plot with quarter-level grouping is presented in Appendix Figure A.8. Standard errors are clustered at the quarter level.

Figure 6: Event Study Plot: Touristic Restaurants Have Relative Improvement in Ratings After Pandemic, Review-Level Specification



Note: This figure displays the results of the event study concerning the pandemic-induced shock, where the unit of analysis is review. Point estimates and 95% confidence intervals are provided. The first quarter of 2020 serves as the excluded time period. Standard errors are clustered at the quarter level.

Figure 7: Matrix of Correlations Between Five Types of Concerns in Reviews Measured as Cosine Distances in *word2vec*



Notes: This figure presents a correlation matrix for five distinct types of concerns found in review texts: (1) mentions of tourists, (2) poor food quality, (3) excessive prices, (4) long waiting times, and (5) excessive noise in the restaurant. Each variable is continuous and measured as the cosine distance between the centroid of words from a review and the centroid of a dictionary related to the subject, using a word2vec embedding trained on the primary corpus of reviews.

Table 1: Tourism Decreases Resident's Satisfaction with Urban Amenities: Restaurant Level Analysis, Difference-in-Differences

Natural experiments:	**Before and After First Pandemic Lockdown** *(Post = Post-Lockdown)*		**Before and After Terrorist Attack – November 2015** *(Post = Post-Terrorist Attack)*			
Dependent variables:	Avg. Rating by Parisians		Avg. Rating by Parisians		Avg. Rating by Non-Parisians	
	(1)	(2)	(3)	(4)	(5)	(6)
Share of Non-French Reviews prior to observation period (by Restaurant) × Post	0.0752***	0.0811***	0.0384***	0.0335***	0.0078	0.0069
	(0.0197)	(0.0238)	(0.0094)	(0.0107)	(0.0090)	(0.0101)
Fixed-effects						
Restaurant	Yes	Yes	Yes	Yes	Yes	Yes
Month	Yes		Yes		Yes	
Month × Neighborhood		Yes		Yes		Yes
Fit statistics						
Observations	75,876	75,876	41,611	41,611	60,309	60,309
R^2	0.35637	0.38035	0.36487	0.38716	0.33306	0.34983
Dependent variable mean	0.71498	0.71498	0.68987	0.68987	0.73798	0.73798
Dependent variable SD	0.3094	0.3094	0.2808	0.2808	0.2255	0.2255

Notes: This table presents OLS estimates. In all columns, the unit of analysis is a month × restaurant pair. The table displays results for two notable shocks in tourism: the first year of the pandemic and the 2015 terrorist attack. The dependent variable in Columns 1-4 is the average rating of restaurants among users with a home location in Paris. For the terrorist attack, we additionally measure the effect on ratings by tourists in Columns 5-6. All ratings are scaled from 0 to 1. The share of non-French reviews is calculated for the periods up to 2019 and up to 2014, corresponding to the pandemic and terrorism-induced shocks, respectively. The *Post* variable is a dummy, activated after each shock. Standard errors clustered at the neighborhood level are in parentheses. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

Table 2: Tourism Decreases Resident's Satisfaction with Urban Amenities: Pandemic Shock, Review Level Analysis, Difference-in-Differences

Dependent variable:	Rating in Review Left by Parisian			
	(1)	(2)	(3)	(4)
Share of Non-French Reviews (by Restaurant) × Post-Lockdown	0.0691*** (0.0209)	0.0470* (0.0240)	0.0656** (0.0298)	0.0847** (0.0389)
Fixed-effects				
Restaurant	Yes	Yes	Yes	Yes
Month	Yes	Yes		
User		Yes	Yes	
Month × Neighborhood			Yes	Yes
User × Post-Lockdown				Yes
Fit statistics				
Observations	120,568	120,568	120,568	120,568
R^2	0.28159	0.73543	0.74618	0.76204
Dependent variable mean	0.71999	0.71999	0.71999	0.71999
Dependent variable SD	0.3312	0.3312	0.3312	0.3312

Notes: This table presents OLS estimates. In all columns, the unit of analysis is an individual review. The dependent variable is the rating of a review left by a Parisian user. All ratings are scaled from 0 to 1. The share of non-French reviews is calculated for the periods up to 2019. The *Post-Lockdown* variable is a dummy variable, activated after the first pandemic lockdown. Standard errors are clustered at the neighborhood level. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

Table 3: Drop in Tourism has Decreased Complaints about Tourists but has not Affected Other Types of Complaints: Pandemic Shock, Review-Level Analysis, Difference-in-Differences

Dependent variables: *(Topics of complaints)*	Tourists	Low Food Quality	Too Expensive	Too Noisy	Long Wait
	(1)	(2)	(3)	(4)	(5)
Panel A: Dictionary-Based (dummy := 1 if contains the keywords from the topic)					
Share of Non-French Reviews	-0.0891***	-0.0032	-0.0334	0.0145	-0.0332
× Post-Lockdown	(0.0222)	(0.0311)	(0.0278)	(0.0265)	(0.0223)
Fixed-effects					
User × Post-Lockdown	Yes	Yes	Yes	Yes	Yes
Restaurant	Yes	Yes	Yes	Yes	Yes
Month × Neighborhood	Yes	Yes	Yes	Yes	Yes
Fit statistics					
Observations	111,756	111,756	111,756	111,756	111,756
R^2	0.56827	0.60988	0.53738	0.47727	0.53808
Dependent variable mean	0.02274	0.07506	0.05095	0.02816	0.02702
Panel B: Word Embedding (Cosine Distances scaled to vary from 0 to 1)					
Share of Non-French Reviews	-0.0209**	-0.0053	-0.0095	-0.0088	-0.0059
× Post-Lockdown	(0.0091)	(0.0141)	(0.0099)	(0.0092)	(0.0101)
Fixed-effects					
User × Post-Lockdown	Yes	Yes	Yes	Yes	Yes
Restaurant	Yes	Yes	Yes	Yes	Yes
Month × Neighborhood	Yes	Yes	Yes	Yes	Yes
Fit statistics					
Observations	111,756	111,756	111,756	111,756	111,756
R^2	0.56827	0.60988	0.53738	0.47727	0.53808
Dependent variable mean	0.02274	0.07506	0.05095	0.02816	0.02702

Notes: This table presents OLS estimates. The unit of analysis is a review. Panels A and B introduce two different measures of concerns expressed in reviews. In Panel A, the dependent variable is derived from the texts of reviews using dictionary-based method. In Panel B, the dependent variable represents the cosine distance between the centroid of words in reviews and the centroid of the dictionary related to a topic in word2vec embedding. The share of non-French reviews is calculated for the periods up to 2019. Post-lockdown is a dummy, which is switched on in June, 2020. Standard errors clustered at the neighborhood level are in parentheses. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

Table 4: Effect of Tourism on Residents' Satisfaction with Amenities Is Stronger for Restaurants Popular Among Tourists from Countries with Weaker Social Ties to France

Dependent variable:	Avg. Rating by Parisians		
	(1)	(2)	(3)
Share of Non-French Reviews × Post-Lockdown	0.0768**		
	(0.0301)		
Share of Non-French Reviews × Post-Lockdown × High Social Connectedness Index		0.0367	0.0406
		(0.0329)	(0.0376)
Share of Non-French Reviews × Post-Lockdown × Low Social Connectedness Index		0.0781***	0.0845***
		(0.0221)	(0.0302)
Fixed-effects			
Restaurant	Yes	Yes	Yes
Month x Neighborhood	Yes		Yes
Month		Yes	
Fit statistics			
Observations	62,001	62,001	62,001
R^2	0.36648	0.33666	0.36652
Dependent variable mean	0.70152	0.70152	0.70152

Notes: This table presents OLS estimates. In all columns, the unit of analysis is the pair: month × restaurant. The dependent variable represents the average rating of restaurants among users with a home location in Paris. All ratings are scaled from 0 to 1. The share of non-French reviews is calculated for the periods up to 2019. The *Post-lockdown* variable is a dummy that is activated in June 2020, following the first COVID-19 lockdown. Measures of network proximity between countries of origin are constructed using Facebook data. Restaurants are categorized based on their proximity score into two groups: those above the median proximity (High SCI) and those below it (Low SCI). Standard-errors clustered at the neighborhood level are in parentheses. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

Table 5: Effect of Tourism on Complaints about Tourists is Higher for Restaurants Popular Among Tourists from Countries with Weaker Social Ties to France

Dependent variables: *(Topics of complaints)*	Tourists (1)	Low Food Quality (2)	Too Expensive (3)	Too Noisy (4)	Long Wait (5)
Share of Non-French Reviews	-0.0491***	0.0197	0.0295	0.0043	-0.0162
× Post-Lockdown	(0.0096)	(0.0177)	(0.0334)	(0.0241)	(0.0130)
× High Social Connectedness Index					
Share of Non-French Reviews	-0.0816***	-0.0221	0.0077	0.0171	-0.0135
× Post-Lockdown	(0.0160)	(0.0247)	(0.0183)	(0.0120)	(0.0135)
× Low Social Connectedness Index					
Fixed-effects					
Restaurant	Yes	Yes	Yes	Yes	Yes
Month × Neighborhood	Yes	Yes	Yes	Yes	Yes
Fit statistics					
Observations	62,079	62,079	62,079	62,079	62,079
R^2	0.24497	0.22017	0.18684	0.18442	0.18753
Dependent variable mean	0.02580	0.07424	0.04878	0.02452	0.02618

Notes: This table presents OLS estimates. In all columns, the unit of analysis is the pair: month × restaurant. The dependent variable is derived from review texts using dictionaries and represents the share of reviews related to one of the corresponding topics, aggregated by restaurant-month. The share of non-French reviews is calculated for the periods up to 2019. The *Post-lockdown* variable is a dummy that activates in June 2020. Measures of network proximity between countries of origin are constructed using Facebook data. Restaurants are categorized based on their proximity scores into two groups: those with scores above the median (High SCI) and those below the median (Low SCI). Standard-errors clustered at the neighborhood level are in parentheses. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

Table 6: No Spillovers: Level of Tourism Around Restaurant Does Not Crowd Out Within Restaurant Effect

Dependent variables:	Avg. Rating by Parisians			
	(1)	(2)	(3)	(4)
Share of Non-French Reviews × Post-Lockdown	0.0763***	0.0697***	0.0774***	0.0694***
	(0.0209)	(0.0252)	(0.0255)	(0.0259)
Touristic Area (<100m) × Post-Lockdown		-0.0349		0.0004
		(0.0378)		(0.0388)
Touristic Area (100m-300m) × Post-Lockdown		0.1021*		0.1139*
		(0.0608)		(0.0664)
Touristic Area (300m-500m) × Post-Lockdown		0.0209		0.0295
		(0.0744)		(0.0857)
Touristic Area (500m-1000m) × Post-Lockdown		-0.0916		0.0204
		(0.0728)		(0.1115)
Fixed-effects				
Restaurant	Yes	Yes	Yes	Yes
Month	Yes	Yes		
Month × Neighborhood			Yes	Yes
Fit statistics				
Observations	63,410	63,410	63,410	63,410
R^2	0.34439	0.34445	0.37327	0.37333
Dependent variable mean	0.70393	0.70393	0.70393	0.70393

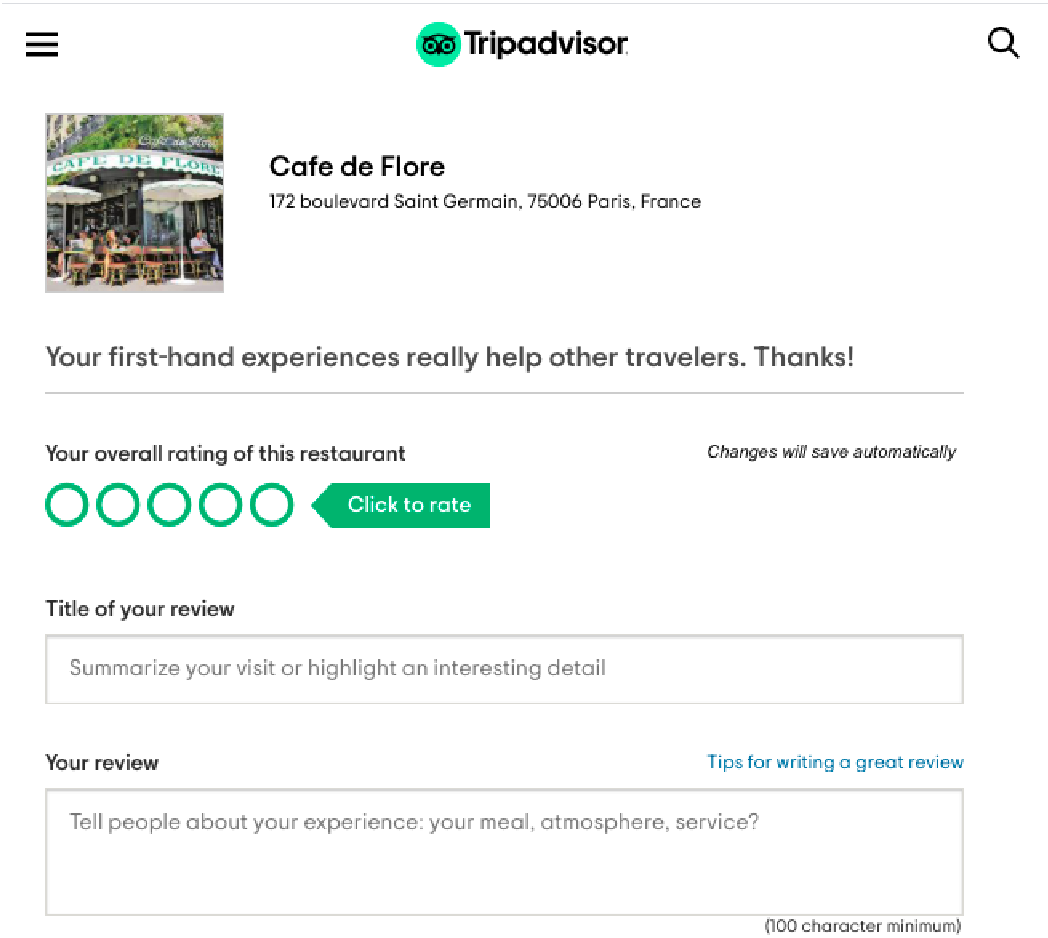
Notes: This table reports OLS estimates. The unit of analysis is a pair month × restaurant. The dependent variable represents the average rating of restaurants given by users whose home location is Paris. All ratings are scaled from 0 to 1. The share of non-French reviews is calculated for the periods up to 2019. The *Post-lockdown* variable is a dummy that activates in June 2020. Standard errors clustered at the neighborhood level. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

Online Appendix

Table of Contents

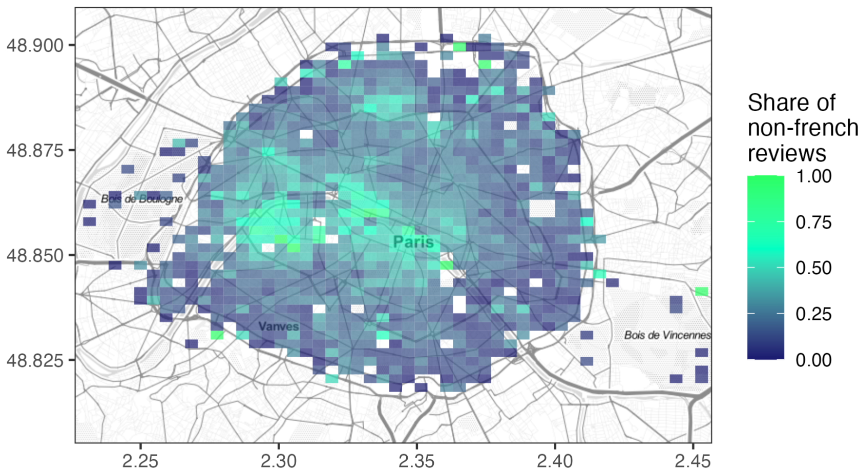
A Additional Plots

Figure A.1: Tripadvisor Interface



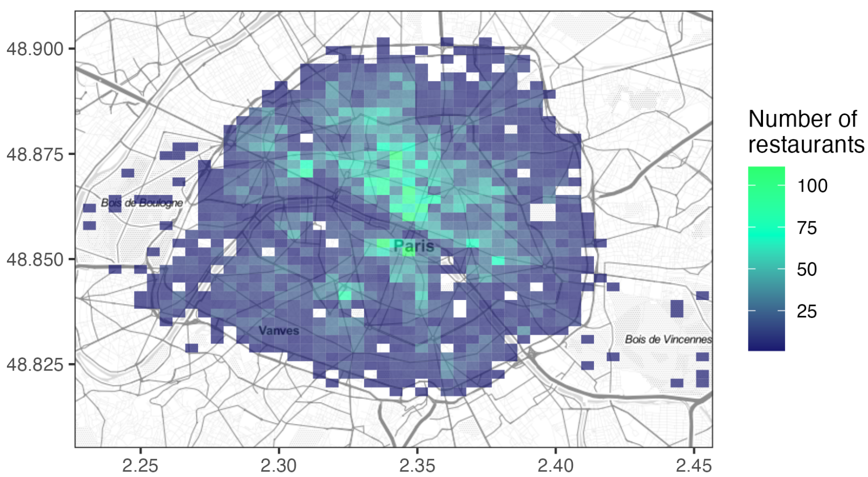
Notes: This figure is a screenshot that users of Tripadvisor see when they want to leave a review. *Café de Flore* is a famous Parisian café-restaurant in the Saint-Germain-des-Prés district, in the 6th arrondissement, popular among tourists and located in close proximity to the Sciences Po campus.

Figure A.2: The Grid Map with the Shares of Non-French Reviews



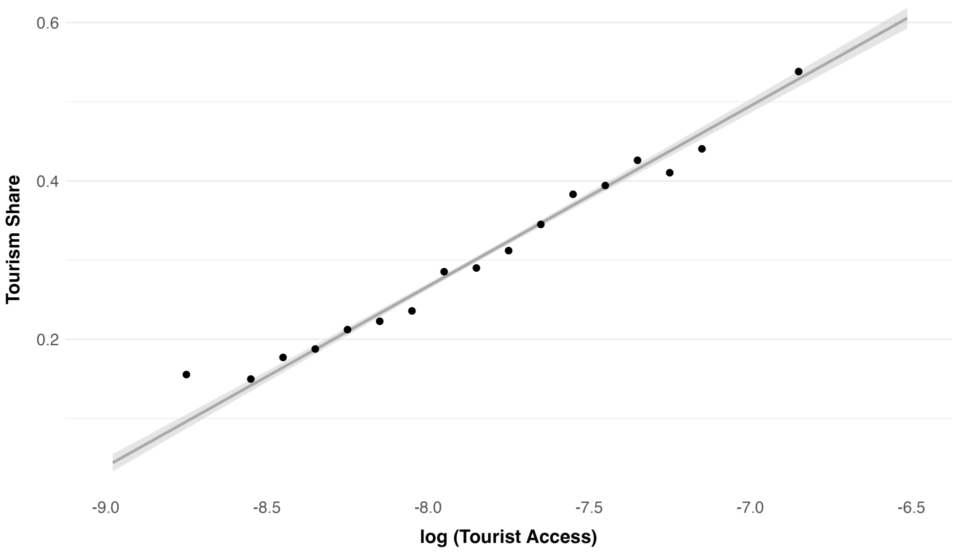
Notes: This map illustrates the distribution of restaurants based on the share of non-French reviews. Each gridcell's color corresponds to the proportion of non-French reviews out of the total reviews prior to 2020 in this area. The color distribution indicates the level of tourist presence in the neighborhood, as the majority of restaurants with a high share of non-French reviews are clustered around major tourist attractions.

Figure A.3: The Grid Map of Restaurant Density Differs from the Measure of Tourism Density



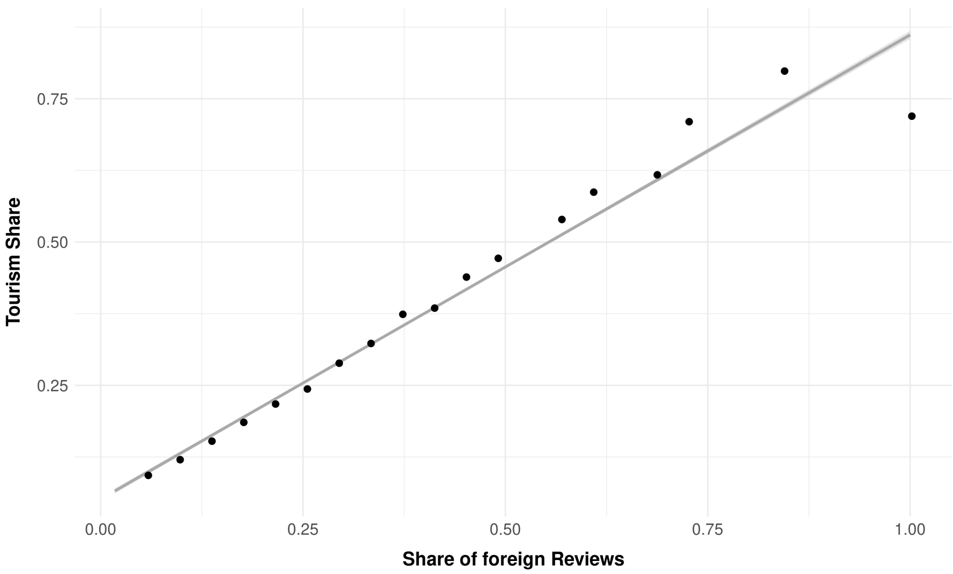
Notes: This map illustrates the density of restaurant locations. Each gridcell's color corresponds to the number restaurants prior to 2020 in this area. This color distribution indicates that the density of restaurants differs from the level of tourist presence in the neighborhood, as illustrated in Figure 3 and Appendix Figure A.2.

Figure A.4: Validation of Tripadvisor Tourism Measure with Tourist Access (Based on Number of Visitors to Major Attractions)



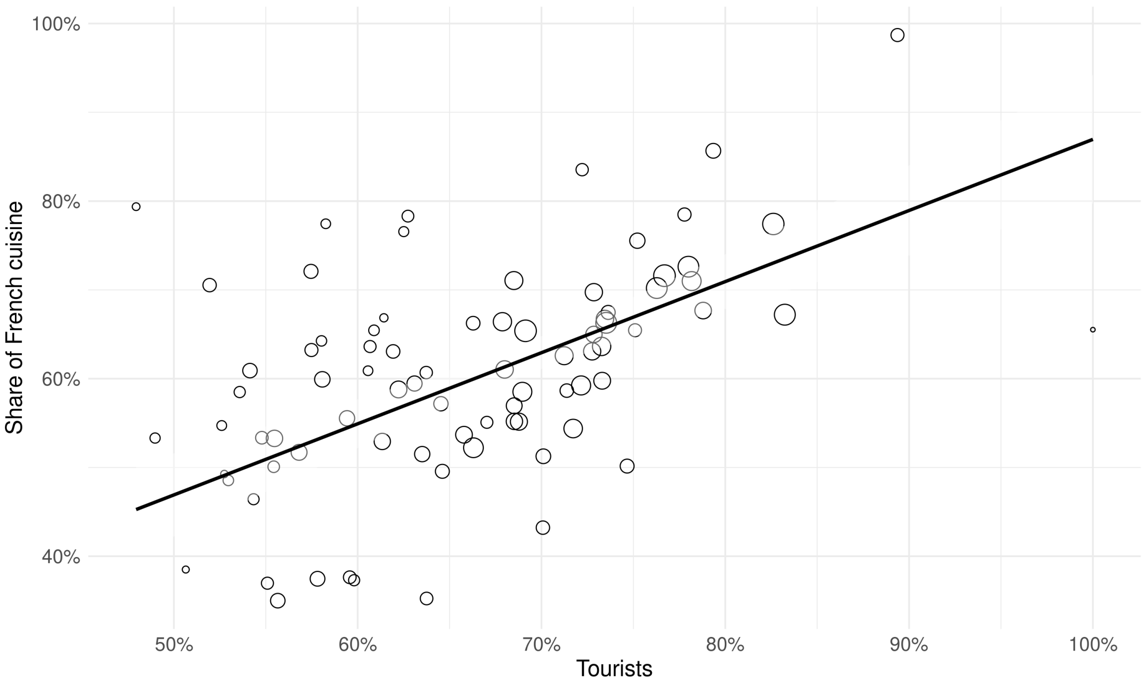
Notes: This binned scatter plot serves as a validation exercise for our baseline Tripadvisor measure of tourism, illustrating its correlation with the presence of tourism as measured by visitor numbers to major tourist attractions. The dataset captures the percentage of tourists in Paris visiting various attractions between 2015 and 2019. We geocoded 18 intra-muros attractions and then constructed a tourist demand measure using the market access framework.

Figure A.5: Correlating Different Tourism Proxies



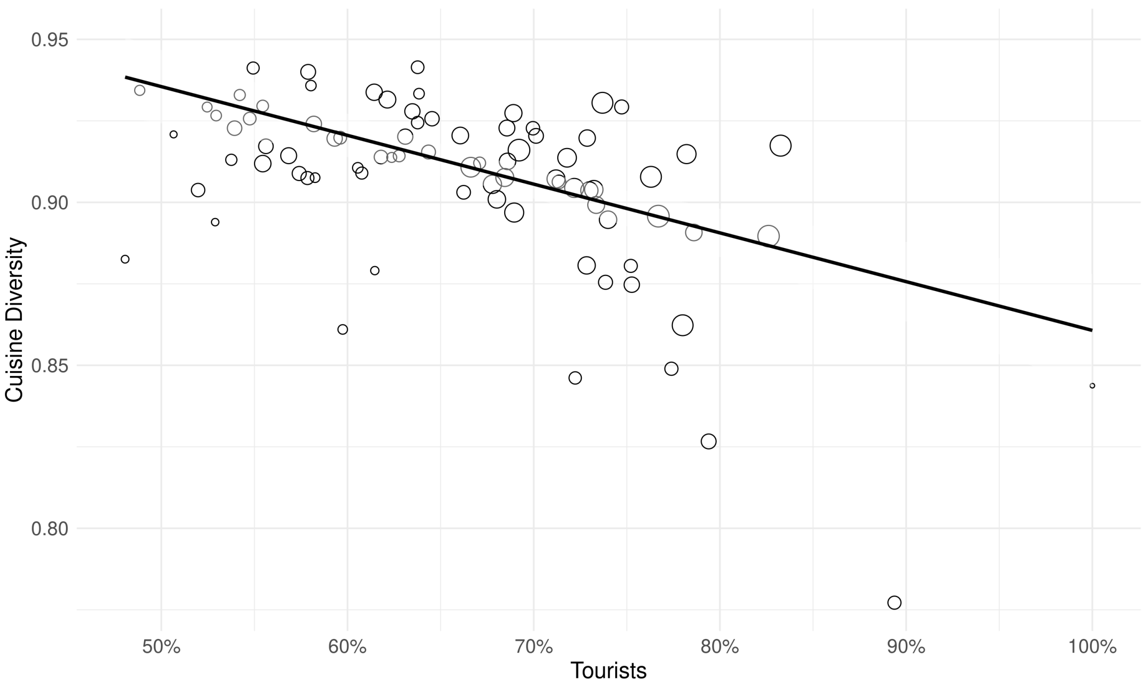
Notes: This figure is a binned scatter plot that illustrates the association between two measures of tourism based on TripAdvisor reviews: the share of non-French reviews and the share of reviews left by users with home location outside of France.

Figure A.6: Share of French Cuisine and Tourism



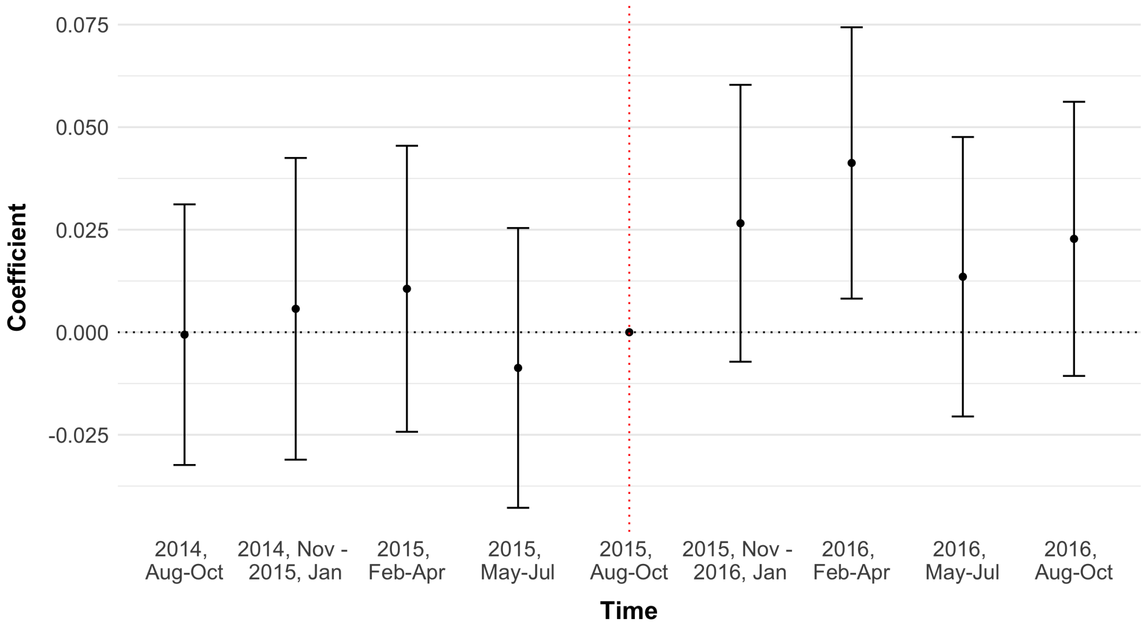
Notes: This figure demonstrates a positive correlation between the language-based measure of tourism and the share of restaurants serving French cuisine, both aggregated at the neighborhood level.

Figure A.7: Diversity of Cuisine Types and Tourism



Notes: This figure shows a negative correlation between the language-based measure of tourism and cuisine diversity, with the latter measured as the reversed Herfindahl-Hirschman Index of cuisine types. Both metrics are aggregated at the neighborhood level.

Figure A.8: Event Study Plot: Touristic Restaurants Have Relative Improvement in Ratings After November 2015 Attack, Restaurant-Level Specification (Quarter Grouping)



Notes: This figure displays the results of the event study concerning the terrorism-induced shock, where the unit of analysis is month × restaurant. Point estimates and 95% confidence intervals are provided. The period from August to October 2015 serves as the excluded time period. Standard errors are clustered at the quarter level.

B Additional Tables

Table B.1: Stylized Facts: Parisian Preferences

Dependent variables:	Rating in Review Left by Parisian		
	(1)	(2)	(3)
Share of Non-French Reviews	-0.0986***	-0.0639***	-0.0774***
(by Restaurant)	(0.0214)	(0.0178)	(0.0175)
log(Number of Reviews)	0.0064*	0.0023	0.0048**
	(0.0032)	(0.0025)	(0.0023)
Fixed-effects			
User		Yes	Yes
Neighborhood			Yes
Fit statistics			
Observations	109,428	109,428	109,428
R^2	0.00276	0.61527	0.61940
Dependent variable mean	0.71658	0.71658	0.71658

Notes: This table reports OLS estimates. In all columns, the unit of analysis is an individual review. Dependent variable is a review's rating by a Parisian. The sample consists of reviews left before the pandemic. The tourism share is measured as the share of non-French reviews left on a restaurant's page up to 2019. Standard errors clustered at the neighborhood level are in parentheses. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

Table B.2: Tourism Decreases Resident's Satisfaction with Urban Amenities: November 2015 Shock, Review Level Analysis, Difference-in-Differences

Dependent variables:	Rating in Review Left by Parisian			
	(5)	(6)	(7)	(8)
Share of Non-French Reviews (by Restaurant) × Post-Attack	0.0277*** (0.0078)	0.0155 (0.0112)	0.0113 (0.0142)	0.0097 (0.0158)
Fixed-effects				
Restaurant	Yes	Yes	Yes	Yes
Month	Yes	Yes		
User		Yes	Yes	
Month × Neighborhood			Yes	Yes
User × Post-Attack				Yes
Fit statistics				
Observations	105,446	105,446	105,446	105,446
R^2	0.23503	0.70385	0.71386	0.75303
Dependent variable mean	0.69882	0.69882	0.69882	0.69882
Dependent variable SD	0.3076	0.3076	0.3076	0.3076

Notes: This table presents OLS estimates. In all columns, the unit of analysis is an individual review. The dependent variable is the rating of a review left by a Parisian user. All ratings are scaled from 0 to 1. The share of non-French reviews is calculated for the periods up to 2014. The *Post-Attack* variable is a dummy variable, activated after the November 2015 terrorist attack. Standard errors, clustered at the neighborhood level, are in parentheses. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

Table B.3: Tourism and "DansMaRue" Complaints

Dependent variables:	# Complaints in "DansMaRue" Application			
	(1)	(2)	(3)	(4)
Share of Non-French Reviews	-0.6570***	-0.2581*		
× Post-Lockdown	(0.2272)	(0.1364)		
Top 25% Most Touristic			-0.3527***	-0.1504**
× Post-Lockdown			(0.1213)	(0.0726)
Fixed-effects				
Restaurant	Yes	Yes	Yes	Yes
Month	Yes		Yes	
Month × Neighborhood		Yes		Yes
Fit statistics				
Observations	366,930	305,332	366,930	305,332
R^2	0.48157	0.68477	0.48024	0.68481
Dependent variable mean	0.40114	0.48207	0.40114	0.48207

Notes: This table reports PPML estimates. The dependent variable is the number of complaints registered on the "Dans ma rue" platform within 100m of a restaurant in a given month. The share of non-French reviews is calculated for the periods up to 2019. *Post-lockdown* is a dummy, which is switched on in June, 2020. Standard errors clustered at neighborhood level are in parentheses. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

C Robustness Checks

C.1 Location-Based Tourism Measure

Table C.1: Location-Based Measure: Tourism and Restaurant Ratings by Parisians: Restaurant-Level Analysis

Dependent variables:	Avg. Rating by Parisians			
	(1)	(2)	(3)	(4)
Share of Non-Parisians	0.1089***	0.0996***		
Reviews × Post-Lockdown	(0.0231)	(0.0246)		
Top 25% Most Touristic			0.0392***	0.0360***
(by Non-Parisians) × Post-Lockdown			(0.0102)	(0.0111)
Fixed-effects				
Restaurant	Yes	Yes	Yes	Yes
Month	Yes		Yes	
Month x Neighborhood		Yes		Yes
Fit statistics				
Observations	75,822	75,822	75,822	75,822
R^2	0.35615	0.38011	0.35608	0.38007
Dependent variable mean	0.71487	0.71487	0.71487	0.71487

Notes: This table presents OLS estimates. In all columns, the unit of analysis is a month × restaurant pair. The dependent variable is the average rating of restaurants among users with a home location in Paris. All ratings are scaled from 0 to 1. The share of non-French reviews is calculated for the periods up to 2019. The *Post-Lockdown* variable is a dummy, activated after the first pandemic lockdown. Standard errors clustered at the neighborhood level are in parentheses. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

Table C.2: Location-Based Measure: Tourism and Restaurant Ratings by Parisians: Review-Level Analysis

Dependent variables:	Rating by Parisian			
	(1)	(2)	(3)	(4)
Share of Non-Parisians Reviews × Post-Lockdown	0.1081*** (0.0245)	0.0807*** (0.0288)	0.0913*** (0.0321)	0.0977*** (0.0356)
Fixed-effects				
Restaurant	Yes	Yes	Yes	Yes
Month	Yes	Yes		
User		Yes	Yes	
Month × Neighborhood			Yes	Yes
User × Post-Lockdown				Yes
Fit statistics				
Observations	120,506	120,506	120,506	120,506
R^2	0.28146	0.73534	0.74611	0.76196
Dependent variable mean	0.71991	0.71991	0.71991	0.71991

Notes: This table presents OLS estimates. In all columns, the unit of analysis is an individual review. The dependent variable is the rating of a review left by a Parisian user. All ratings are scaled from 0 to 1. The share of non-French reviews is calculated for the periods up to 2019. The *Post-Lockdown* variable is a dummy variable, activated after the first pandemic lockdown. Standard errors, clustered at the neighborhood level, are in parentheses. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

Table C.3: Location-Based Measure: Textual Outcomes

Dependent variables:	Tourists (1)	Low Food Quality (2)	Too Expensive (3)	Too Noisy (4)	Long Wait (5)
Share of Non-Parisians Reviews × Post-Lockdown	-0.0562***	-0.0213	0.0013	-0.0014	-0.0165
	(0.0111)	(0.0186)	(0.0155)	(0.0109)	(0.0119)
Fixed-effects					
Restaurant	Yes	Yes	Yes	Yes	Yes
Month × Neighborhood	Yes	Yes	Yes	Yes	Yes
Fit statistics					
Observations	75,943	75,943	75,943	75,943	75,943
R^2	0.24864	0.23044	0.19964	0.18781	0.19802
Dependent variable mean	0.02308	0.07171	0.04730	0.02367	0.02563

Notes: This table presents OLS estimates. In all columns, the unit of analysis is the pair: month × restaurant. The dependent variable is derived from review texts using dictionaries and represents the share of reviews related to one of the corresponding topics, aggregated by restaurant-month. The share of non-Parisian reviews is calculated for the periods up to 2019. The *Post-lockdown* variable is a dummy that activates in June 2020. Standard errors clustered at the neighborhood level are in parentheses. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

C.2 Aggregation of Language-Based Tourism Measure by Different Periods

Table C.4: Tourism and Ratings: Language-Based Tourism Aggregated by Different Periods

Dependent variables:	Avg. Rating by Parisians				
	(1)	(2)	(3)	(4)	(5)
Share of Non-French Reviews (< 2017) × Post-Lockdown	0.0665** (0.0278)				
Share of Non-French Reviews (< 2018) × Post-Lockdown		0.0793*** (0.0270)			
Share of Non-French Reviews (< 2019) × Post-Lockdown			0.0863*** (0.0247)		
Share of Non-French Reviews (< 2020) × Post-Lockdown				0.0811*** (0.0254)	
Share of Non-French Reviews (< 2021) × Post-Lockdown					0.0823*** (0.0274)
Fixed-effects					
Restaurant	Yes	Yes	Yes	Yes	Yes
Month x Neighborhood	Yes	Yes	Yes	Yes	Yes
Fit statistics					
Observations	57,292	65,515	72,112	75,876	76,350
R^2	0.37559	0.37228	0.37469	0.38035	0.38273
Dependent variable mean	0.69755	0.70390	0.71083	0.71498	0.71565

Notes: This table presents OLS estimates. In all columns, the unit of analysis is a month × restaurant pair. The dependent is the average rating of restaurants among users with a home location in Paris. All ratings are scaled from 0 to 1. The share of non-French reviews is calculated for the periods up to a specified year. The *Post-lockdown* variable is a dummy, activated after the first pandemic lockdown. Standard errors clustered at the neighborhood level are in parentheses. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

C.3 Clustering

Table C.5: Tourism and Ratings: Different Clustering

Dependent variables:	Avg. Rating by Parisians		
	(1)	(2)	(3)
Share of Non-French Reviews × Post-Lockdown	0.0811***	0.0811***	0.0811***
	(0.0254)	(0.0245)	(0.0238)
Fixed-effects			
Restaurant	Yes	Yes	Yes
Month x Neighborhood	Yes	Yes	Yes
Clustering			
	Quarter	Grid cell	No
Fit statistics			
Observations	75,876	75,876	75,876
R^2	0.38035	0.38035	0.38035
Dependent variable mean	0.71498	0.71498	0.71498

Notes: This table presents OLS estimates. In all columns, the unit of analysis is a month × restaurant pair. The dependent is the average rating of restaurants among users with a home location in Paris. All ratings are scaled from 0 to 1. The share of non-French reviews is calculated for the periods up to 2019. The *Post-lockdown* variable is a dummy, activated after the first pandemic lockdown. Standard errors clustered at a specified level are in parentheses. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

D Validation of Tourism Measures

Table D.1: Tourist Access

	Tourism Share			
	(1)	(2)	(3)	(4)
log(Tourist Access)	0.2443***	0.2170***	0.2450***	0.1409***
	(0.0171)	(0.0369)	(0.0215)	(0.0326)
Weighted			Yes	Yes
Fixed-effects				
Neighborhood		Yes		Yes
Fit statistics				
Observations	10,179	10,179	10,179	10,179
R^2	0.22746	0.31021	0.26590	0.39319
Dependent variable mean	0.31451	0.31451	0.31451	0.31451

Notes: This table presents OLS estimates. It demonstrates a validation exercise for our baseline measure of tourism, illustrating its correlation with the presence of tourism as measured by visitor numbers to major tourist attractions. The dataset captures the percentage of tourists in Paris visiting various attractions between 2015 and 2019. Standard errors clustered at the neighborhood level are in parentheses. Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

E Text Analysis

Table E.1: Dictionary 1: Words, Phrases and Wildcard Expressions for Dictionary-Based Labelling of Reviews

Low Food Quality

pas bon *(not tasty)*, sans goût *(no taste)*, aucun saveur *(no taste)*, réchauff *(reheated)*, pas très bon *(not very tasty)*, aucun goût *(no taste)*, cuisine bof *(kitchen yuck)*, mauvaise cuisson *(poor cooking)*, goût bizzare *(wierd taste)*, industriel *(industrial)*, avarié *(rotten)*, pas assez cuit *(undercooked)*, trop cuit *(overcooked)*, supermarch *(supermarket)*, tombé malade *(got sick)*, pas cuit *(not cooked)*, sans saveur *(no taste)*, vomir *(vomit)*, mauvaise qualité *(bad quality)*, indigestion *(indigestion)*, intoxication *(intoxication)*, pas frais *(not fresh)*, fade *(no taste)*, surgel *(frozen food)*, insipid *(no taste)*, dégueulass *(disgusting)*, micro-ond *(microwaved)*, pas fait maison *(not "homemade")*

Too Expensive

prix élevés *(high price)*, cher *(expensive)*, prix sont élevés *(high price)*, prix sont très élevés *(very high price)*

Too Noisy

bruyant *(noisy)*, beaucoup de bruit *(a lot of noise)*

Long Wait

long *(long)*, lent *(long)*

Tourism

touris *(tourist/tourism)*

Table E.2: Dictionary 2: Seeding Words for Word Embedding Labelling of Reviews

Low Food Quality

insipide *(no taste)*, indigestion *(indigestion)*, surgelé *(frozen food)*, industriel *(industrial)*, fade *(no taste)*

Too Expensive

cher *(expensive)*, élevé *(increased)*, excessif *(excessive)*, coûteux *(expensive)*, onéreux *(expensive)*

Too Noisy

bruyant *(noisy)*, bruit *(noise)*

Long Wait

long *(long)*, lent *(long)*, tardif *(late)*

Tourism

touriste *(tourist)*, tourisme *(tourism)*

Table E.3: Summary Statistics for Textual Variables (Dictionary-Based)

Variable	N	Mean	St. Dev.
Tourism	1,154,860	0.025	0.157
Low Food Quality	1,154,860	0.066	0.248
Too Expensive	1,154,860	0.050	0.218
Too Noisy	1,154,860	0.028	0.165
Long Wait	1,154,860	0.024	0.153

Table E.4: Ratings and Textual Variables (Dictionary-Based)

Dependent Variable:	Rating by Parisian				
	(1)	(2)	(3)	(4)	(5)
Variables					
$\mathbb{1}\{\text{tourists}\}$	-0.0817***				
	(0.0095)				
$\mathbb{1}\{\text{poor food}\}$		-0.2918***			
		(0.0050)			
$\mathbb{1}\{\text{expensive}\}$			-0.1116***		
			(0.0057)		
$\mathbb{1}\{\text{noisy}\}$				-0.0544***	
				(0.0068)	
$\mathbb{1}\{\text{long wait}\}$					-0.1003***
					(0.0064)
Fixed-effects					
User	Yes	Yes	Yes	Yes	Yes
Restaurant	Yes	Yes	Yes	Yes	Yes
Date	Yes	Yes	Yes	Yes	Yes
Fit statistics					
Observations	121,808	121,808	121,808	121,808	121,808
R^2	0.74061	0.76171	0.74253	0.74033	0.74119
Dependent variable mean	0.72131	0.72131	0.72131	0.72131	0.72131

Notes: This table presents OLS estimates. In all columns, the unit of analysis is an individual review. The dependent variable is the rating of a review left by a Parisian user. All ratings are scaled from 0 to 1. The explanatory variables are dummies that indicate whether the text of a review relates to the given topic (dictionary-based). Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

Table E.5: Ratings and Textual Variables (Word Embedding)

Dependent Variable:	Rating by Parisian				
	(1)	(2)	(3)	(4)	(5)
Variables					
Tourism (cosine dist.)	-1.119***				
	(0.0239)				
Poor food (cosine dist.)		-1.276***			
		(0.0132)			
Expensive (cosine dist.)			-0.6427***		
			(0.0201)		
Noisy (cosine dist.)				-0.8997***	
				(0.0180)	
Long wait (cosine dist.)					-1.211***
					(0.0226)
Fixed-effects					
User	Yes	Yes	Yes	Yes	Yes
Restaurant	Yes	Yes	Yes	Yes	Yes
Date	Yes	Yes	Yes	Yes	Yes
Fit statistics					
Observations	113,193	113,193	113,193	113,193	113,193
R^2	0.76520	0.81321	0.75322	0.76197	0.76656
Dependent variable mean	0.72160	0.72160	0.72160	0.72160	0.72160

Notes: This table presents OLS estimates. In all columns, the unit of analysis is an individual review. The dependent variable is the rating of a review left by a Parisian user. All ratings are scaled from 0 to 1. The explanatory variables are cosine distances to the centroids corresponding to the given topic (word embedding). Signif. Codes: ***: 0.01, **: 0.05, *: 0.1