ISSN 2282-6483

Alma Mater Studiorum - Università di Bologna
DEPARTMENT OF ECONOMICS

Tourist Tax and Ratings of Online Reviews

Antonio Marsi
Emanuela Randon

Quaderni - Working Paper DSE N°1168

Tourist Tax and Ratings of Online Reviews *

Antonio Marsi[1] and Emanuela Randon[2]

[1]CAST and Department of Economics, University of Bologna

[2]Department of Economics, University of Bologna

December 13, 2021

Abstract

Tourist taxes represent one of the principal sources of revenue for a number of local authorities. Using a dataset of more than 300 thousands reviews on TripAdvisor, we analyze the effect of tourist taxes on online ratings posted by hotel costumers. Online ratings are strategic variables for the competitiveness of hotels and local destinations and so the assessment of the impact of tourist taxes on online reviews is important to design suitable strategies and policies both for firms and local municipalities. We show that the share of complaints about the tax increases with the tax rate, but the relationship gets weaker as the hotel quality increases, suggesting that low-quality costumers are more sensitive to tax increases. Tax mentioning reviews have on average a 20% lower rating and this difference becomes smaller and not significant for high quality hotels. We disentangle the different sources of complaining, finding that the lack of information and the cash payments weight more negatively on the average rating (respectively less 27% and 21%). Using a random-effects logit model we show that what matters for costumers is the percentage tax on the room price and not the absolute amount of the legal tax rates based on the hotel stars. This effect is present only for hotel stays with an average double room price under 100 Euros. Our results provide public authorities with useful suggestions to change the actual tax system by setting the tourist tax as an ad valorem tax on room price and assessing optimal tax rates accordingly to their incidence on the price of the stay.

Keywords: Tourism Tax; Online Ratings; Tax Incidence.

JEL Codes: D12,H21,H22,L83,Z3.

*We are grateful to Patrizia Battilani, Flavio Delbono, Paolo Figini, Marco Magnani and to all the participants of the 2021 CAST-Days for useful comments and suggestions. We gratefully acknowledge funding by Uni.Rimini.

Non technical summary

We investigate the relationship between the tourist tax and online reviews by looking at a sample of online reviews about hotels located in a set of famous touristic destinations in Italy. The tourist tax is paid by tourists on overnight stays at hotels and it is set according to a per-person-per-night rule. It constitutes an important source of revenues for Italian municipalities.

It is natural to ask whether this tax produces some negative side effect on the satisfaction of paying tourists. To answer this question we analyze a sample of more than 300 thousands online reviews posted by tourists on TripAdvisor. By exploiting variations in the tax rates across cities and time, we are able to draw some important results.

First, we show that a fraction of reviews are actually mentioning the tourist tax, and that these reviews are associated to a lower rating on average. The fraction of these tax-mentioning reviews is positively correlated with the tax rate. Moreover, the degree of dissatisfaction embedded in these reviews (measured by the difference in their ratings with respect to the total average) is higher in low-quality hotels (with a lower number of stars), suggesting an heterogeneity in the sensitivity of tourists to the presence of the tax.

Second, we compare the ratings of all the online reviews to the tax rate, using a random-effects logit model, and find that on-average online ratings do not vary across different levels of the legal tax rate, which is a per-person-per-night amount and vary across hotel categories. If instead we construct an effective tax rate measure, i.e. the fraction of the tax on the average price of the hotel, we identify a significant negative correlation between online ratings and the effective tax rate. This results suggest that the employment of an ad valorem tax system should be favored by policy makers, as it allows to avoid distortions created by the actual specific tax system.

1 Introduction

The recent decades have seen a formidable increase in tourism around the world due to the fast process of globalization. In the EU tourism represented 9.1% of the GDP in 2019 and it is expected to increase its economic importance, despite the temporary drastic fall due to the pandemic in 2020. In Italy the average yearly increase in tourist arrivals over the years 2000-2014 has been almost 2% (3% for foreign tourists), which implies that arrivals in 2014 were 34% higher than in 1999, 54% if we consider only non-residents. This fast increase in the demand for tourism has prompted a surge in taxes on tourism activities, (European Commission, 2017) which can take several forms: VAT, visa fees, entry/exit charges, taxes on hotels and restaurants charged by local or central governments. The accommodation tax is the most popular form of tourist taxation for local municipalities[1] (Gooroochurn and Sinclair, 2005) and it is recognized as an important source of local public financing for two main reasons. Firstly, it is a politically convenient instrument of local taxation since the welfare losses are born by taxpayers that do not vote for local governments (Gooroochurn and Sinclair, 2005). Secondly, it contributes to internalize the negative effects of the externalities on the resident population and on the environment due to over-tourism (Gooroochurn and Sinclair, 2005). Though the rationales for taxing tourists are clear and extensively analyzed in the literature, little research has been made on the perception of tourist taxes by customers and on how this perception is related to the incidence of taxation. So far, the analysis of the incidence of the tourist tax is linked to the estimation of the aggregate price elasticity of tourism demand, by looking at the number of tourists' arrivals (Candela *et al.*, 2013; Aguiló *et al.*, 2005; Biagi *et al.*, 2017). For instance, Aguiló *et al.* (2005) estimates a drop of 1.44% of daily arrivals of foreign tourists in the Balearic Islands in response to 1 Euro tax increase. In this paper, we analyze the effect of tourist tax on the rating of online reviews in order to study how customers perceive the tourist tax and whether this perception influences the overall satisfaction about the quality of the hotel experience. Our findings suggest effective strategies and policy interventions both to firms' managers and to local authorities to improve the competitiveness of hotels and local destinations.

The role of online reviews in consumers' decision making process has grown rapidly in the last decade, spanning a wide range of products and services (Chevalier and Mayzlin, 2006; Vermeulen and Seegers, 2009). The tourism sector is as well largely involved in this process,

[1] Italy, for example, collected a total of 600 millions Euros of revenues in 2019, while this figure amounted to 500 millions Euros in France.

with a number of websites where consumers can look for accommodations, browse online reviews made by previous customers, book a room and leave reviews after the stay. Keeping a high reputation on these websites is a paramount aspect of the business strategies of hotels. Chen *et al.* (2016) have shown that a higher review quality, measured by the average ratings posted on TripAdvisor, significantly increases the economic performance of hotels, measured by their occupancy ratio. On the opposite side, a lower average rating damages hotels' performance. Thus the analysis of the impact of taxes on the ratings of online reviews is important both for the firms operating in the tourist sector and for policy makers. If higher taxes negatively affect consumers' online ratings, hotels are inevitably damaged and can lose competitiveness. This possible effect of tourist taxes is thus worth studying to suggest strategies of optimal pricing to hotel owners that should consider the impact of pass-through effects.[2] Moreover, local tourism taxes affect the overall competitiveness and attractiveness of tourist destinations. Identifying the perception that tourists have on such taxes may help policy makers to better design tax rates in order to balance the need to raise revenues for public finances and the need to maintain the competitiveness of the local destination (Schubert, 2010; Agrawal *et al.*, 2021; Aguiló *et al.*, 2005; Chiarello, 2013; Besley and Case, 1995).

Occupancy taxes typically are set per person on overnight stays and are collected by hotels or similar categories. In Italy the tax has been introduced in 2011, allowing municipalities to decide whether to enforce it and to what level set tax rates. Only the city of Rome was allowed to introduce the tax earlier than 2011. In Italian municipalities, tax rates are set according to two main principles: the tax should be progressive and tax revenues should be used by municipalities to invest in local infrastructures and services for environmental purposes or to finance their touristic supply. Furthermore, several exemptions apply and differ among municipalities. Usually the tax is applied only to the first seven nights of the stay and children are exempted. To implement the desired progressiveness, tax amounts vary according to the category of the hotel, i.e. the number of stars.[3]

To study whether the tax jeopardizes the satisfaction of tourists we look at online ratings

[2]The existing literature suggests a high-degree of pass-through of tourist taxes on the price paid by hotel customers (see European Commission, 2017, for a review of studies analyzing the impact of tourist taxes on tourism in Europe), i.e. a 1 euro increase in the occupancy tax raises the gross price paid by tourists by a similar amount.

[3]This system is the most common across European countries, with few exceptions (e.g. some regions of Germany, Hungary, Romania) where the tax is set as percentage of the price (European Commission, 2017).

posted on TripAdvisor. After staying at a hotel, tourists are allowed to review their experience and the quality of the hotel. This is carried out by mean of a textual review and a rating (from one to five) assigned at the hotel. This rating is affected by the price of the room(Wolf *et al.*, 2020), the quality of the service received and by the overall opinion of the tourists about their experience at the hotel.

We analyze a sample of approximately 300.000 online reviews from TripAdvisor which have been scraped in the beginning of 2020. This sample covers the whole history of all the hotels, located in five Italian cities, reviewed in TripAdvisor from 2011 to 2019.[4] We first compute the overall distribution of ratings, noticing that, as usual, is extremely skewed towards the maximum value: 46% of the posted ratings are equal to five. The overall average is 4.14. Our sample also confirms the usual result that the posted average rating for each hotel category is increasing with the number of stars, giving evidence of the positive correlation among ratings and hotel quality. We then move to the analysis of the impact of the occupancy tax on online reviews considering different perspectives. First, we focus on reviews that mention the tax and we show that their rating is about 20% lower than the average rating of the other reviews. We disentangle three main different reasons for complaining about the tax in these reviews: the excessive amount of the tax (24% of the posted reviews), the lack of information about the existence of the tax till the payment (25%) and use of cash for its payment (6%). The complaining that more negatively affect the rating is the lack of information about the tax: it lowers the score by almost 27% (from 4.14 to 3.04), followed by the cash (less 21% - from 4.14 to 3.25) and the size (less 18%- from 4.14 to 3.38). Moreover we show that as we move to higher quality hotels, the frequency of tax-mentioning reviews increases with the tax rate, but at a decreasing rate, implying that customers of high-price hotels increase their complaints to tax increase less strongly than customers of low-price hotels. This suggests that low-quality costumers are more sensitive to tax increases than high-quality hotels.

In a second step, we analyze the full set of reviews and study whether ratings differ depending on the size of the tourist tax applied for the stay at the hotel. This analysis is implemented by mean of a random-effects logit regression model. Costumers are almost indifferent to the absolute value of the tax (i.e. the legal value in euros for each hotel category per night staying): the probability of having a five rating is constant whatever is the size of the legal tax in euros. On the contrary, costumers are quite sensitive to the weight of the tax on the price (i.e. the percentage tax rate): we show that a 10% increase in the incidence of the

[4]We skipped hotels which display less than twenty reviews on total.

tax on the price reduces the probability of having a 5 rating by 4%. The effect is larger for mid-star hotels, where the probability of having a 5 rating decreases by 6%, while it is absent for 5-star hotels. In detail, we show that the negative impact of the percentage tax increase is significant only for hotels with an average price for a double room smaller than 100 Euros. As a complementary result, we find a positive correlation among tax rates and low ratings, ranged between 1 and 4. These results are found using a random-effects panel logit model; in the Appendix we show that these results are confirmed also using random-effects ordered logit regressions.

The rest of the paper proceeds as follows. After a section with an overview of the literature (Section 2), we describe the dataset of online reviews gathered from TripAdvisor (Section 3). In Section 4 we look at tax-mentioning reviews and study the relationship between the amount of tax rates, the frequency of tax-mentioning reviews and ratings, analyzing their linkages and effects by hotel category. In Section 5 we run a random-effects logit regression for panel data on the full sample of online ratings to see whether costumers are sensitive to the legal tax on the stars or to the ad valorem tax on the price. Policy suggestions conclude the paper.

2 Review of the literature

Since tourist taxes are an important source of revenue for local public finance, our paper is strictly connected to all the literature on policy choices of local governments, that has been recently reviewed by Agrawal *et al.* (2021). Furthermore, our paper relates to: (i) the growing number of studies analyzing the rationale for taxing tourists and the effect of these taxes on the economy; (ii) the vast literature analysis the determinants of online ratings, in particular regarding the tourism sector and to iii) the literature on specific versus ad valorem taxes.

The 2017 Report by the European Commission extensively analyzed the role of tourist taxes, in their various forms, in the European Union. The report highlights the fact that compared to other taxes such as VAT the accommodation tax can have additional negative effects on tourists, as it is generally not observable in the published accommodation prices and it is usually paid at the departure. It is more seen as an "extra" rather than a tax which consumers are used to account for. In this regard, it is paramount to understand how the tax is perceived by tourists. For this reason some studies have estimated, through surveys,

factors influencing the willingness to pay and tourists' awareness for a tax of this kind (see Do Valle *et al.*, 2012; Duràn-Romàn *et al.*, 2021; Rodella *et al.*, 2019; Borges *et al.*, 2020). Borges *et al.* (2020) found that approximately 50% of respondents were aware of the tax through a survey run in Porto (Portugal). The final incidence effect has been analyzed in a few papers. For example, Bonham and Gangnes (1996) found that a tourist tax introduced in Hawaii was entirely shifted onto consumers. The impact found on hotel revenues is however minimal, suggesting that the demand is quite inelastic. Aguiló *et al.* (2005) estimate the elasticity of the demand for tourism from foreign tourists in the Balearic Islands, and quantify the drop in tourist arrivals induced by an increase in the tax, concluding that a tax of 1 Euro per tourist per day would lead to a decrease in annual visitor arrivals of 117,000 to the region. Candela *et al.* (2013) analyze the introduction of the tourist tax in some tourist destinations in Italy in the province of Rimini and estimate that a relevant reduction in tourism flows is observed in Riccione due to substitution effects towards a neighboring municipality (i.e. Misano), whereas no effect emerges in the case of two adjoining jurisdictions (i.e. Rimini and Cattolica). Biagi *et al.* (2017) estimates the impact of the introduction of the tourist tax in a municipality in Sardinia (i.e. Villasimius). They found a small decline in domestic demand for tourism and no effect for international tourism. These aspects make it challenging for municipalities to optimally determine the amount of the tax, as they have to maximize tax revenues while trying to mitigate negative side-effects (Schubert, 2010). In this regard, online ratings provide a unique data set for investigating the perception of taxes among costumers. The role of consumers' reviews on online platforms and social media in various markets for goods and services has received a strong attention in the economic literature in the last period, due to the growing importance of this type of channel for conveying information among consumers. Vermeulen and Seegers (2009) show that for hotels, exposure to online reviews improves their performance and positive reviews have major beneficial impact. Chen *et al.* (2016) show that an higher review quality, measured by the average ratings posted on TripAdvisor, significantly increases the economic performance of the hotels, measured by their occupancy ratio. On the opposite side, a lower average rating damages hotels' performance. Xiang and Gretzel (2010) highlights the growing importance played by online platforms in the online search made by tourists when planning their travels. As online ratings have become a crucial component in the accommodation sector, a large number of studies have focused on what factors influence online ratings most. Besides obvious factors related to the quality of the service provided by hotels (see Ho *et al.*, 2020), the extraordinary amount

of data has allowed researchers to focus on possible determinants which are usually not taken into account. See, for instance, Gao *et al.* (2018), that look at how cultural factors such as "power distance" are reflected in ratings posted on TripAdvisor. Finally, Wolf *et al.* (2020) study the relationship between consumers' satisfaction and the price paid for the service, suggesting a "money for the value effect": everything else equal, online ratings decrease with the price.

To the best of our knowledge no paper has yet investigated the impact of taxes on online ratings. Moreover, the occupancy tax constitutes an interesting example of a specific tax, as the amount to be paid by tourists only depend on the quantity of the service consumed (the number of overnight stays) and only indirectly on the price, as the tax is allowed to vary across hotel categories. Our paper is thus linked to the literautre that compares the allocative effects of specific vs ad valorem taxes. For Example, Denicolò and Matteuzzi (2000) shows that ad valorem taxes are superior in a monopolistic competition setting where firms compete à la Cournot. Wang *et al.* (2018) shows that this could not be the case if firms are heterogeneous in both their cost functions and in the quality of the service offered. Our analysis shows clearly that even if it is a specific tax, it is the degree to which the tax weights on the price of the stay that matters for driving consumers (dis)satisfaction, suggesting new strategies for the optimal setting of tax rates by local authorities.

City	Hotel 1 star	Hotel 2 stars	Hotel 3 stars	Hotel 4 stars	Hotel 5 stars	Total
Roma	12	57	192	177	29	467
Milano	7	19	34	27	8	102
Firenze	6	19	46	24	7	95
Napoli	2	1	13	18	0	34
Rimini	6	36	134	19	0	195
Total	33	132	419	265	44	893

Table 1: Number of hotels observed for each category in each city.

3 Data

We scraped all the reviews for accommodation facilities[5] posted on TripAdvisor from 2011 to 2019 for five Italian cities: Rome, Milan, Naples, Florence and Rimini, gathering a data set of approximately 400.000 reviews[6]. We chose these cities since they are located in the first places in the ranking of Italian cities that collect most revenues from the tourist tax. In order to address our research question, we firstly matched each review in the sample to the level of the tourist tax applied at the time of the visit reported in the review. We observed the year and month of the visit, therefore by knowing the category of the facility, we could merge our sample with official data from the municipalities (see Table 2 for a record of tourist tax rates over time). This procedure was possible since TripAdvisor reports both the type of facility (i.e. an hotel or a Bed and Breakfast) and the number of stars of the hotel. The former information is self-declared, while the latter is provided to TripAdvisor from a third party. An issue arise since many of the facilities which identify as B&Bs on TripAdvisor are often hotels, according to the categorization implemented by the Italian law. For this reason we re-categorized B&Bs which contain either the word "hotel" or "albergo" in their name as hotels, dropping all remaining B&Bs. We are left with a dataset of 315.869 reviews from 893 different hotels. Table 1 summarize the number of hotels observed for each category in each city. More than 50% of the hotels observed are in Rome followed by 22% of hotels in Rimini. Almost half of the hotels are 3-stars and almost 30% are 4-stars.

Since 2011, when the national legislation allowed municipalities to levy the tax, these cities have implemented and adjusted it in different periods of time. Furthermore, legal tax rates have been increased several times over the period considered. Table 2 summarizes all variations in the tax imposed on a per person-per night basis in the different cities. This variation in the legal tax rate across time and cities allows us to implement panel data techniques to analyze the effects of change in the tax itself (see Section 5).

Even if the total number of reviews observed for each hotel varies substantially across hotels, the average number of reviews observed for each hotel is quite high (354 reviews). As Figure 1 shows, the distribution of the number of reviews for each hotel is positively skewed with one fourth of the observed hotels having less than 100 posts and with few hotels displaying more than 1000 posts.

4-stars and 5-stars hotels are usually bigger and display an higher number of posted reviews

[5] We skipped accommodation facilities with less than 20 reviews on total.

[6] Reviews are written in Italian or translated automatically in Italian by TripAdvisor.

Month	City	Hotel 1 star	Hotel 2 stars	Hotel 3 stars	Hotel 4 stars	Hotel 5 stars
Jan2011	Roma	0 » 2	0 » 2	0 » 2	0 » 3	0 » 3
Jul2011	Firenze	0 » 1	0 » 2	0 » 3	0 » 4	0 » 5
Jul2012	Napoli	0 » 0	0 » 1	0 » 2	0 »3	0 » 4
Sep2012	Milano	0 » 1	0 » 2	0 » 3	0 » 4	0 » 5
Oct2012	Rimini	0 » 0.5	0 » 0.7	0 » 1.5	0 » 2.5	0 » 3
Jul2013	Milano	1 » 2	2 » 3	3 » 4	4 » 5	5 » 5
Oct2013	Napoli	0 » 1	1 » 1	2 » 1.5	3 »2.5	4 » 4
Jul2014	Roma	2 » 3	2 »3	2 » 4	3 » 6	3 » 7
Apr2015	Firenze	1 » 1.5	2 »2.5	3 »3.5	4 » 4.5	5 » 5
Aug2015	Napoli	1 » 1	1 » 1.5	1.5 » 2	2.5 » 3	4 » 4
Apr2017	Napoli	1 » 1.5	1.5 » 2	2 » 2.5	3 » 3.5	4 » 4.5
Jan2018	Firenze	1.5 » 2	2.5 » 3	3.5 » 4	4.5 » 4.8	5 » 5
Jan2018	Milano	2 » 2	3 » 3	4 » 4	5 » 5	5 » 5
Jan2019	Rimini	0.5 » 0.7	0.7 » 1	1.5 » 2	2.5 » 3	3 » 4
Jul2019	Napoli	1.5 » 1.5	2 » 2	2.5 » 3	3.5 » 4	4.5 » 4.5

Table 2: Changes in the legal tax rate, measured in Euros per person per night, across cities.

(reviews of 4-stars hotels constitutes 50% of our sample, though 4-star hotels are only 30% of the total number of hotels). We will deal with the issue of over-representativeness in our empirical analysis, by selecting a random subsample of reviews for each hotel.

For each review the online user is posting a rating from 1 to 5, which measures his/her satisfaction with the stay at the hotel. The overall average rating is 4.14. As usual in online ratings, the distribution of the ratings is extremely skewed towards the highest values (see Figure 2): 46% of the posted ratings are equal to 5 and more than 30% are equal to 4. As we can see from the left panel of Figure 3, the average rating increases with the number of stars of the hotel, as the average quality of the hotel increases. Moreover, the highest increase in the rating is between 4 and 5 stars, while the variation in rating between 3 and 4 stars is almost null.

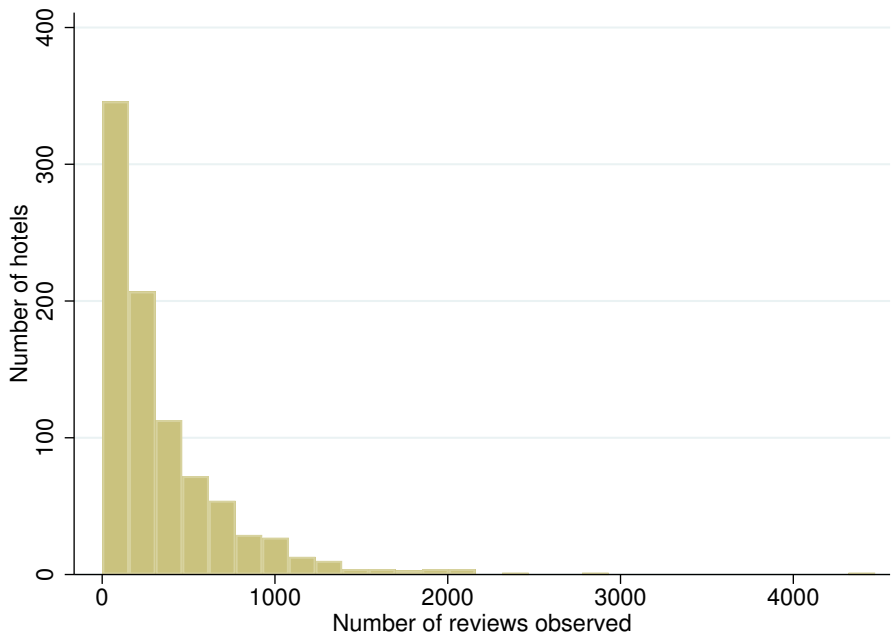

Figure 1: The distribution of the total number of reviews observed across hotels in the sample.

3.1 Legal tax vs effective tax

For each hotel we also observe the price-range reported by TripAdvisor, which indicates the average price-range of a standard-double room. The right panel of Figure 3 shows how the reported average prices varies across hotel categories. There is obviously an increasing pattern, which is convex with respect to the number of stars, reporting a sizable increase from 4-star to 5-star hotels and a flat pattern between 1 and 2 stars. Using this information (which is missing for 9% of the observed hotels) we compute the effective % tax rate as the percentage ratio of the tax amount (per person per night) over the middle value of the price range (dividing by two as it refers to a double room). Figure 4 compares the distribution - across hotel categories - of the legal tax rates (per person per night, on the left side) and the effective tax rates, on the right hand side. The constructed effective tax rates mimic an ad valorem system. As expected, official tax rates drastically increase on average with the stars of the hotels, confirming the progressive nature of the legal tax rate. In details, the highest increase in the legal tax rate is registered between 3 and 4 stars, followed by 1 and 2 stars. It is quite interesting, however, the observation of how the effective tax rates varies with the numbers of stars. This aggregate pattern does not display any clear progressive behavior, with the sound result that the median value of the % tax rate of 5-star hotels is the lowest one. In Figure 5 we disaggregate the pattern looking on how the average tax

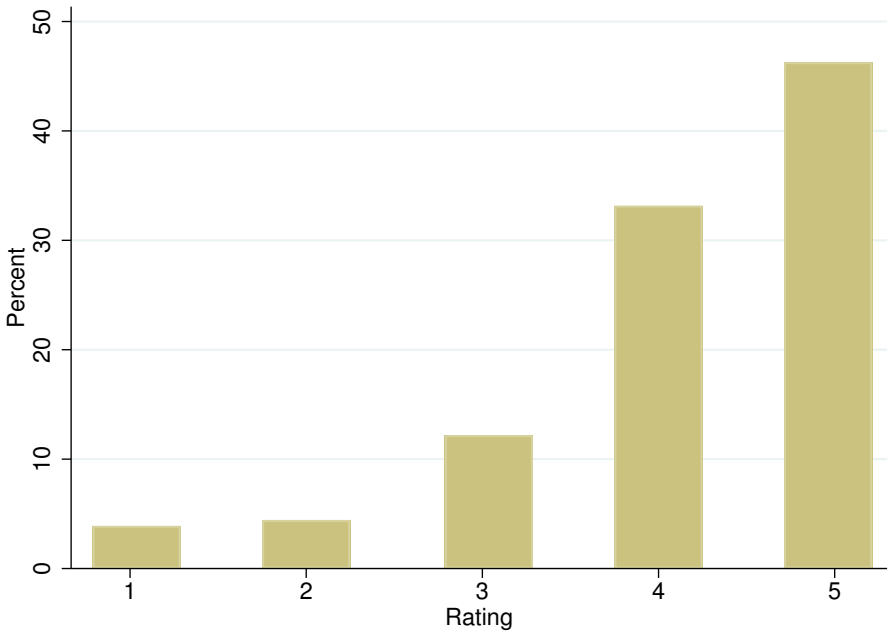

Figure 2: Distribution of ratings' value.

rates for each hotel category vary in each city in the sample, considering only 2018 data. Interestingly, a clear progressive pattern of the % tax rate is observed only for Rimini, while a U-shaped pattern emerges for most of the other cities. Throughout the paper, we will compare the legal tax rates and the effective ones (% tax rates) to assess the incidence and the distributional effect of the tourist tax on costumers.

The number of posted reviews is varying across the years, following a usage of TripAdvisor which is varying over time. Most of the reviews (75%) have been posted between 2013 and 2017. [7] On the contrary, reviews are quite evenly spread across the months of the year, with a slightly higher share referring to the summer period.

[7] In details, this is how the sample is spread across the years: 3% of reviews have been posted in 2011, 9% in 2012, 12% in 2013, 14% in 2014, 17% 2015, 19% in 2016, 13% in 2017, 6% in 2018, 5% in 2019.

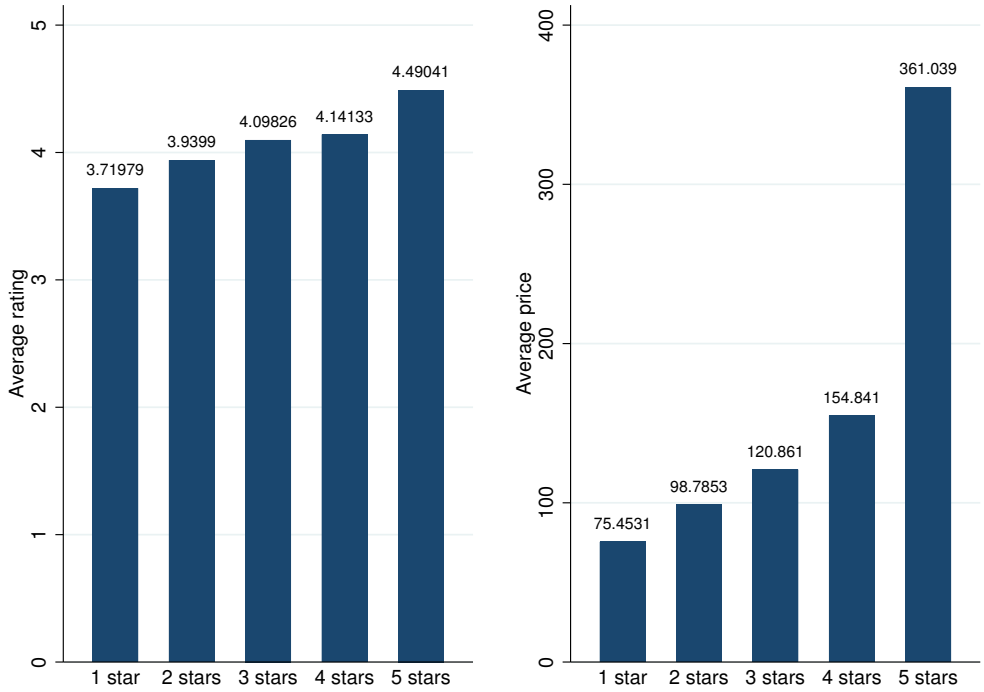

Figure 3: Average rating (left panel) and average price for a standard double room (right panel) for each hotel category.

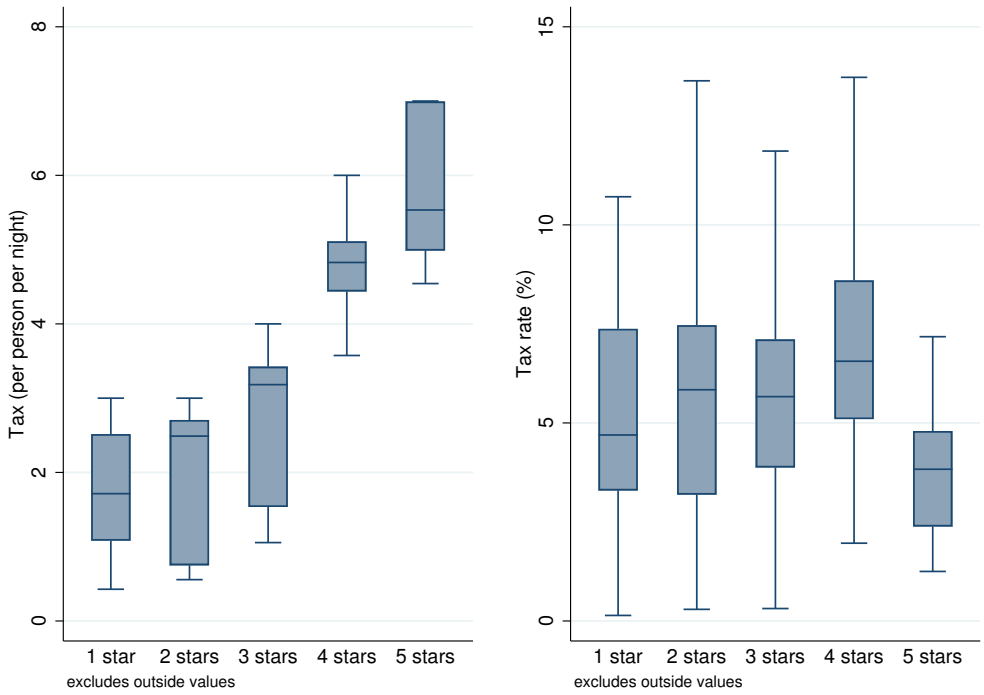

Figure 4: Median % tax rate and 25th,75th percentiles for each hotel category.

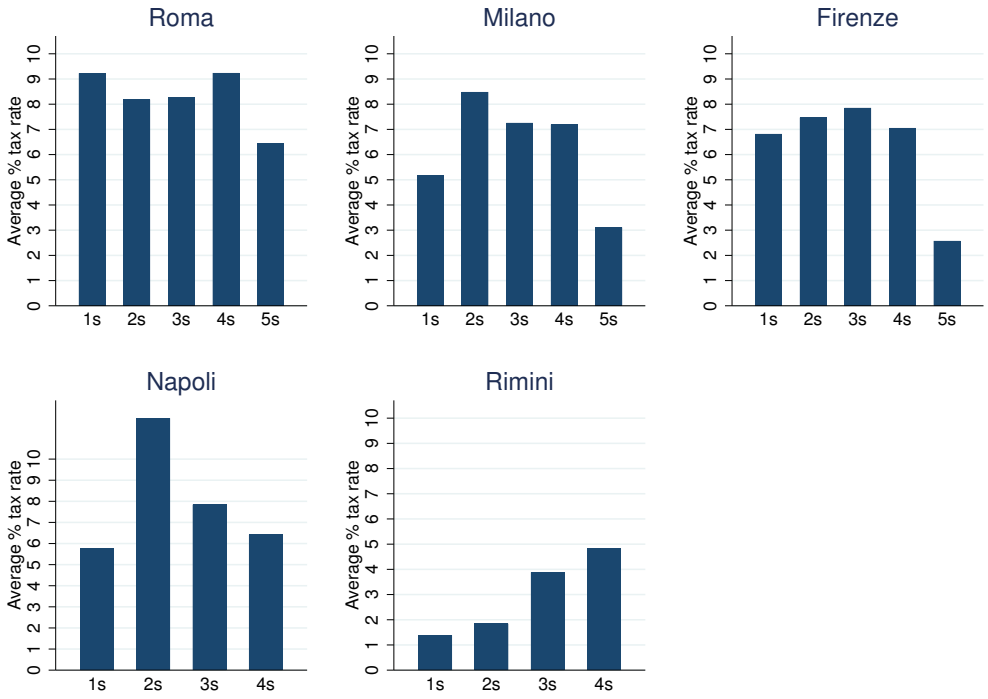

Figure 5: Average % tax rate for each hotel category in each city in 2018.

4 Tax-mentioning reviews

In this section, we focus on tax-mentioning reviews and we start looking at how the probability of mentioning the tax in the review is related to the tax amount and the quality of the hotel. We then analyze the impact of the occupancy tax on the average online rating and on the average rating for each quality category of hotels. In order to identify the "tax-mentioning reviews" we create a dummy variable that is equal to 1 for all the reviews containing the word "tax" ("*tassa/e*", "*imposta/e*" in Italian) and zero for the others. Overall 0.76% of the reviews are tax-mentioning.

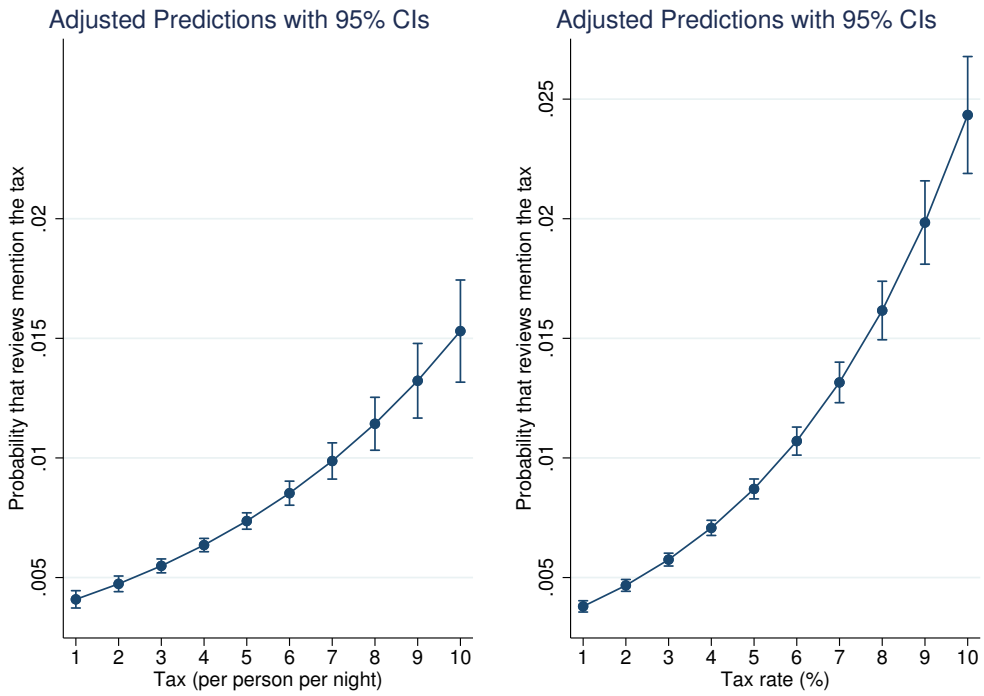

Figure 6: Estimated probabilities that reviews mention the tax (logistic regression).

In details, we firstly show that the probability of a review mentioning the tax is positively correlated with the amount of the tax itself. We run a Logit regression using the dummy for tax-mentioning reviews as dependent variable versus the tax rate linked to that review. We use both the legal tax rate and the % tax rate. Figure 6 depicts how the estimated probability varies with the tax rate, reporting the results of the legal tax rates on the left hand side and the results of the % tax rates on the right hand side. In both specifications we observe a clear increasing pattern, which is more pronounced when using the % tax rate. We then run the same logistic regression, using the % tax rate, distinguishing for each hotel

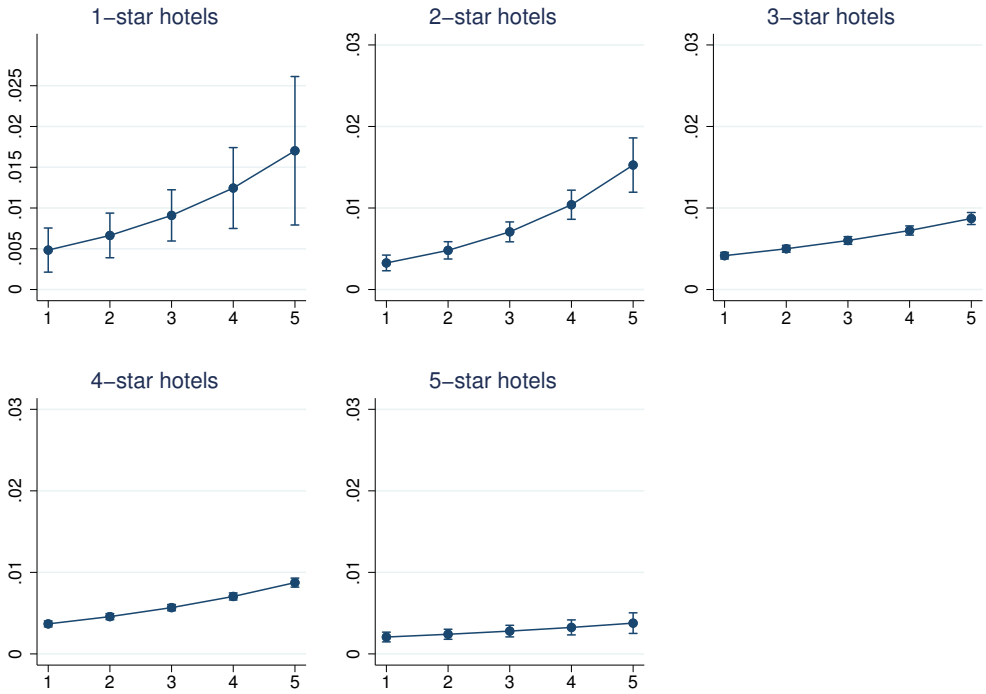

Figure 7: Estimated probabilities that reviews mention the tax vs the tax rate (logistic regression run separately for each hotel category).

category. Figure 7 displays how the predicted probability of mentioning the tax varies with tax rates. The slope of this relationship is clearly decreasing as we move into higher-quality categories, suggesting that customers of high-quality hotels are responding less strongly to increases in the tax rate. These results suggest that costumers from low quality hotels are more sensitive to tax increases than high-quality costumers.

We then look at how mentioning the tax in the review is reflected on the online ratings, finding that the rating of the reviews mentioning the tax is on average almost 20% lower than the rating of the remaining reviews (from 4.14 to 3.34, see Figure 8). We go further in the investigation, distinguishing the different reasons of complaining and their incidence on the ratings. We read all the reviews mentioning the tax (which are approximately 2000) and disentangle three reasons of complaining: the high amount paid for the tax (24% of tax mentioning reviews), the use of cash as tool of payments instead of ATM/credit cards (6%), the lack of information about the existence of the tax that is revealed only at the time of payment (25%) [8]. The strongest negative effect on the rating is due to the lack of information about the tax, which lowers the rating by 27% (from 4.14 to 3.04), followed by

[8]The remaining tax reviews are simply mentioning the tax without complaining about it.

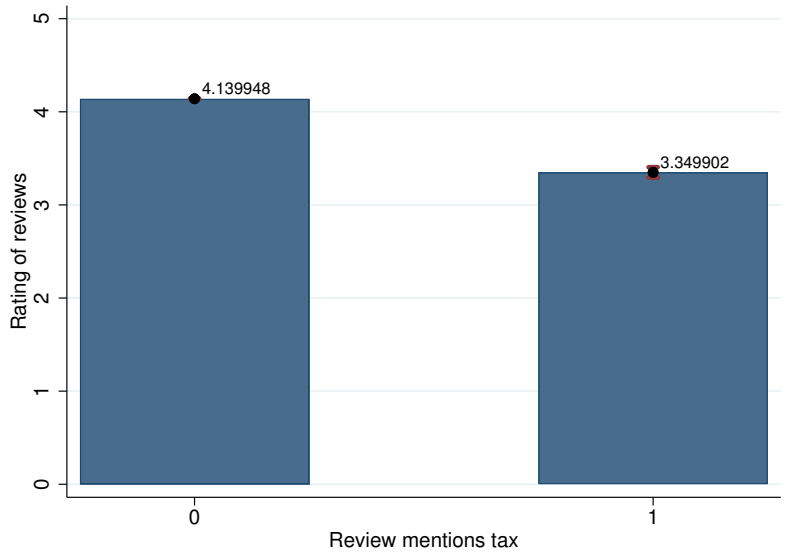

Figure 8: Average rating of reviews not mentioning the tax (0) and mentioning the tax (1).

the use of cash, that records a 21% decrease (i.e. from 4.14 to 3.25%) and the amount of the tax (less 18% , from 4.14 to 3.38). However also the subsample of the reviews that simply reports the tax without complaining about it registers a 16% lower score.

We then look at how the magnitude of the discontent embedded in tax-mentioning reviews varies across different hotel categories. This magnitude can be inferred from the difference between the average rating of tax-mentioning reviews versus the overall average rating. As we can see from Table 3 the negative effect of mentioning the tax on the rating is decreasing in the number of stars[9], and it is mainly concentrated on 2-star hotels (-1.137 in ratings) and 3-star hotels (-0.825). This confirms a decreasing negative effect of the tax as we move into high-quality (the negative effect of the tax on the ratings for 4 and 5-star hotels is below the average, respectively -0.70 and -0.672).

These results allow us to draw some preliminary conclusions about how the tourist tax negatively affect the experience of tourists. The share of tourists complaining for the tax is relatively higher in low-quality hotels. This result is intuitive, as these low-income consumers are likely more watchful about their expenses. An unexpected tax has thus significant effects on their valuations of their stays. Nonetheless, the tax on these hotels is also lower, though evidently not lower enough to compensate the higher attention devoted to the price paid for the accommodation. When taking into account the extent to which these complaints are reflected in a lower rating of the reviews, most significant negative effects are found in 2 and

[9]With the exception of 1-star hotels, where the coefficient is not significant.

	ALL	Hotel 1 star	Hotel 2 stars	Hotel 3 stars	Hotel 4 stars	Hotel 5 stars
Tax-mentioning	-0.790***	-0.524*	-1.137***	-0.825***	-0.701***	-0.672***
	(0.0281)	(0.2215)	(0.1127)	(0.0488)	(0.0367)	(0.1513)
NT	315869	3997	20584	110181	156658	24449

Table 3: Difference in average ratings of tax-mentioning reviews with respect to overall average, for full sample and for each category separately. * p<0.05, ** p<0.01, *** p<0.001. Robust standard errors in parenthesis.

3-star hotels (Table 3).

5 Tourist tax and ratings of reviews

In the previous section, we have shown results on the linkages between tax mentioning reviews, the amount of the tax and the average rating overall and by category. Though these results are interesting and shades some light on the reaction of online reviewers to higher tax rates, it refers only to a small percentage of the total number of reviews. In other words, the effect found could be limited to a small percentage of tourists which are both uninformed about the tax or very attentive about their expenses. A natural question to ask then is whether the negative effect of the tax on reviews could be also found when taking into account the whole sample of reviews, indifferently from whether the tax is mentioned or not in the review. In order to tackle this issue, we would ideally compare two hotels which are perfectly equal in terms of all their characteristics but that only differ with respect to the tourist tax. If the hotel with higher tax was displaying a significantly lower average rating in the reviews received, then the tax effect on online reviews would have been disentangled. To answer this question we run a random effects logit model of the ratings on the tax rate. Following the standard methodology used when using online ratings, we treat the rating variable as a dummy variable that is equal to 1 if the rating is 5 and 0 if the rating is less than 5. This transformation of our dependent variable mirrors the main variability of the rating variable, as its distribution is extremely skewed towards the 5 value (see Figure 2). In Appendix A, we perform a robustness check of our methodology, repeating the analysis on the full range of possible values of ratings through the implementation of an ordered-logit model: we show that results are unchanged. In order to avoid biases in the representativeness

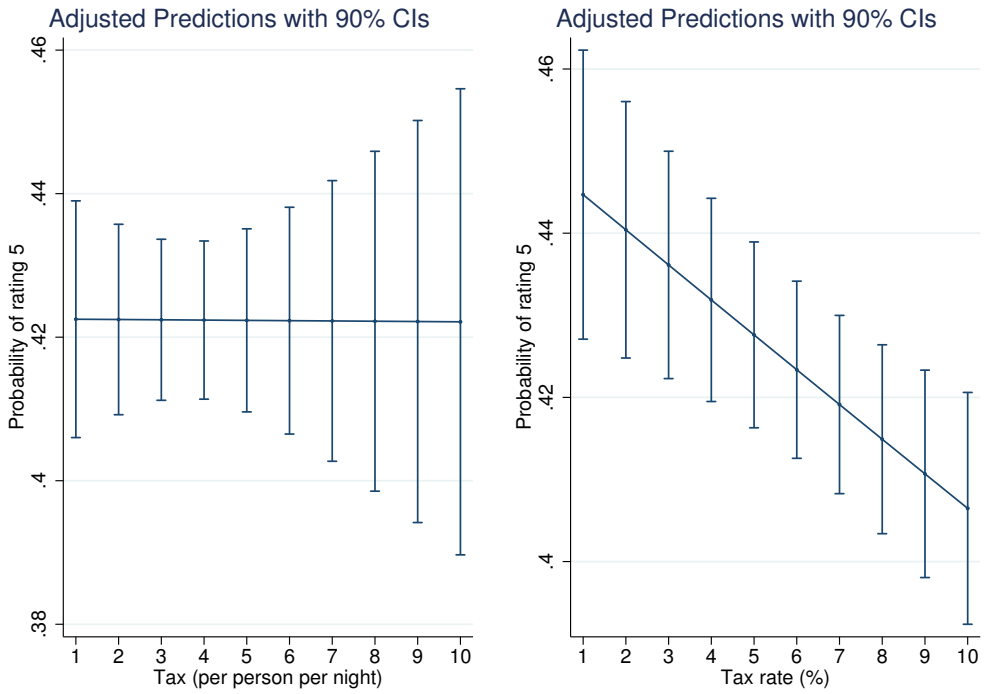

Figure 9: Estimated probabilities of ratings being equal to 5 from the logistic regression on the tax rate (legal tax rate, per person per night, on the left, % tax rate on the right). Vertical bars show 90% confidence intervals.

of the sample, in the following analysis we randomly select, for each hotel, a maximum of 200 reviews, thus shrinking the sample size to 135,260 observations. This random subsampling procedure helps not only to have a more balanced panel but also to ease the computational burden of the estimation required for the random-effects-logit algorithm. Without this selection, the sample would give too much weight to the few hotels with a lot of reviews, inducing a sample selection bias. This bias arises because in these hotels it is high the probability that managers invite reviewing costumers expected to have a good opinion of the hotel stay, inflating the rating score.

The logit model estimates the probability of observing a positive outcome (Rating is equal to 5), denoted by $p_{h,i}$, for the ith reviews belonging to hotel h:

$$p_{h,i} = \frac{e^{y^*_{h,i}}}{1 + e^{y^*_{h,i}}} \tag{1}$$

where $y^*_{h,i}$ is the dependent variable of a linear model:

$$y^*_{h,i} = \alpha + \beta TaxRate_{h,i} + \pi X_{i,h} + u_h + \varepsilon_{i,t} \tag{2}$$

$$\pi X_{i,h} = \psi Price_h + \sum_j \gamma_j DType_h + \sum_k \delta_k DYear_{h,i} + \sum_m \tau_m DMonth_{h,i} + \sum_l \phi_l DCity_h \tag{3}$$

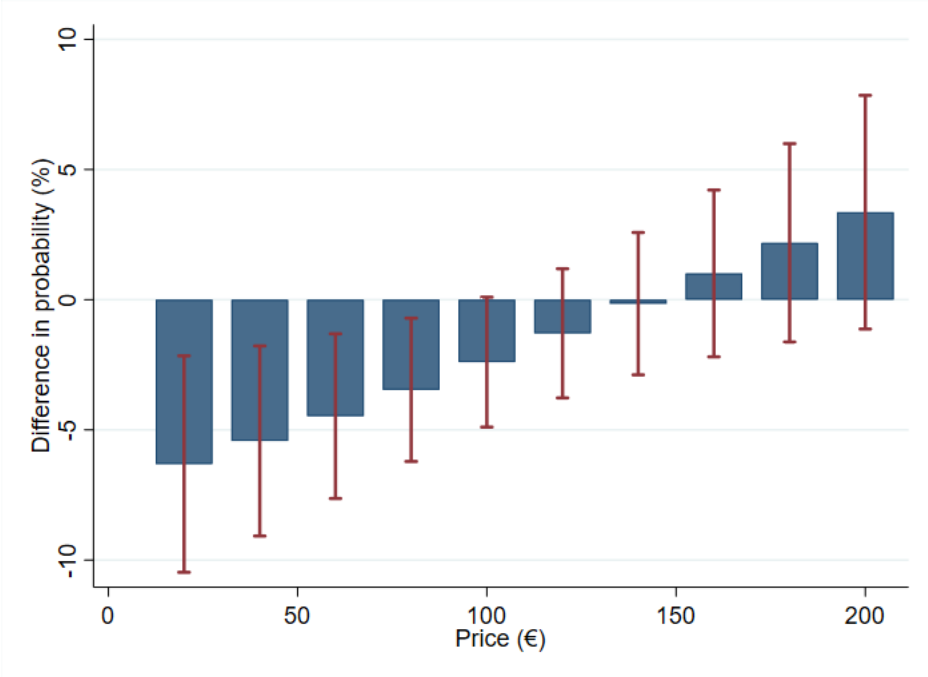

Figure 10: On the y-axis the estimated difference in probability $\hat{P}(Rating = 5/Tax\ rate = 10\%) - \hat{P}(Rating = 5/Tax\ rate = 1\%)$ is depicted, for different values of the price of a standard room. Vertical red bars denote 90% confidence intervals.

We study two scenarios: firstly we consider the case in which the explanatory variable of interest is the legal tax rate, and secondly we assess the analysis on the % tax rate. In both scenarios, we use the same set of control variables: the average price of the hotel ($Price_h$), the year (using the dummy variable $DYear_{h,i}$) and the month ($DMonth_{i,h}$) in which the reviewer has visited the hotel, the city of the hotel ($DCity_h$) and the star-category to which the hotel belongs ($DType_h$).

Our identification strategy relies on the assumption that β is constant across cities. This allows us to have enough variation in the tax across time and cities, which would not be the case if we were focusing on single cities. In fact most of them have changed the legal tax rate only once or two in the time span of the sample (see Table 2).

In Appendix B, we show the full list of estimated coefficients from the logit regression, while here we focus only on our main variable of interest: the tax rate. Figure 9 displays how the estimated probability of posting a rating equal to 5 is varying with tax rates in the two scenarios. Vertical bars show 90% confidence intervals. On the left hand it is assessed the case of the legal tax rate, which in the graph is allowed to range from 1 Euro to 10 Euros. In this case, the relationship between the estimated probability of posting a rating equal to 5 and the tax amount is completely flat. Once the % tax rate is considered, instead, it emerges

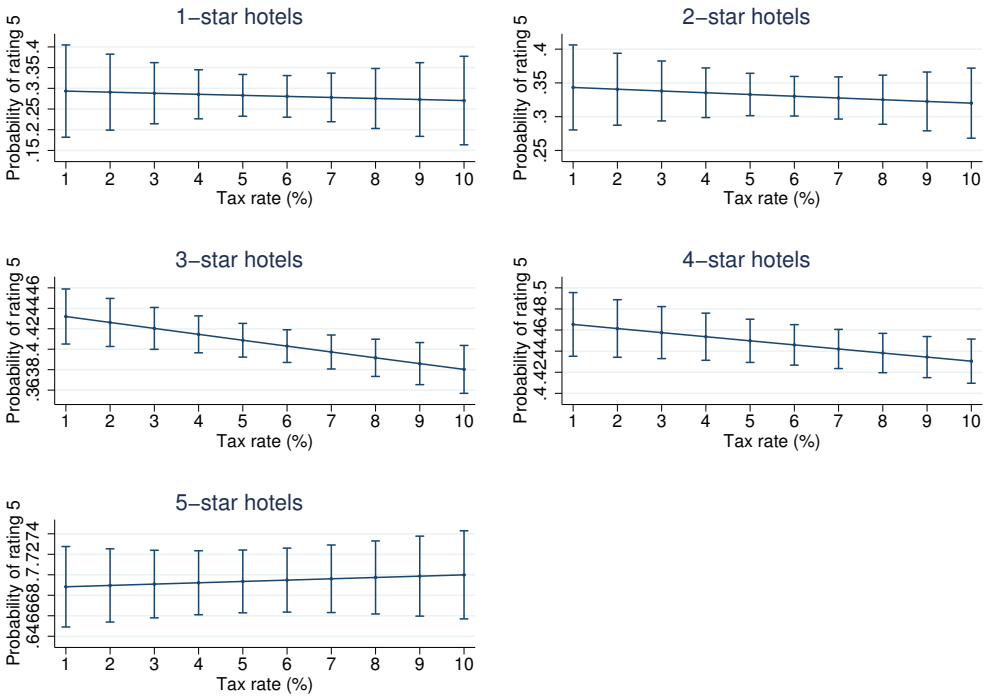

Figure 11: Estimated probabilities of rating being equal to 5 versus the tax rate, across hotel categories. Vertical bars denote 90% confidence intervals.

a clear negative pattern: a tax rate increasing from 1% to 10% reduces the probability of posting a 5-rating of about 4%. This evidence states clearly that what matters for tourists' satisfaction is the degree to which the tax weights on the price, not its absolute value.

It is interesting to understand whether the estimated negative effect varies across the quality of the hotel observed. We expect customers of cheap hotels to be more sensitive to increases in the tax. To verify this hypothesis we run the logit model in equations (1)-(3) by adding an interaction term between the % tax rate and the observed average price of the hotel, thus letting the β coefficient in Equation (2) to vary with the price. We obtain the results shown in Figure 10, where we can see the estimated difference between the probability of posting a rating 5 when the tax rate is 10% and the probability of posting a rating 5 when the tax rate is 1%. This estimated difference in probabilities is displayed for several price ranges, reported on the x-axis. These results confirm our hypothesis: the negative effect produced by the higher tax rate is concentrated in the hotels with an average price (for a double room) lower than 100 Euros, while the effect is not significant for more expensive hotels. These results provides further evidence on the fact that customers reacting more strongly to tax rates are those visiting low-price hotels.

Figure 11 shows instead how the estimated relationship varies across hotel categories. A significant negative effect is found only for 3-star hotels where the probability of having a 5 rating decreases by 6%, while it is absent for 5-star hotels confirming the previous results that costumers of high quality hotels are insensitive to the current level of occupancy taxes.

6 Conclusion

We have analyzed the effect of occupancy taxes on tourists' assessments about the quality of their stays at hotels, by exploring the impact of different tax rates on online ratings posted in TripAdvisor. Our data set contains more than 300 thousands reviews, covering the whole history of online posts on hotels of five major Italian cities (Rome, Milan, Naples, Florence and Rimini), from the introduction of the tax in 2011 until 2019. Tax rates changed frequently in the period and exploiting this variation across and within hotels allow us to use a random-effects panel logit model.

We firstly focused on tax mentioning reviews. Even if these reviews are increasing with the tax rates, the slope of this relationship is decreasing with hotel stars, suggesting a weaker reaction of customers of high-quality hotels to increases in the tax rate. We go further in the analysis and show that tax-mentioning reviews display a 20% lower rating that the other reviews. The most negative effect is concentrated on 3-star hotels and it is absent in 5-star hotels. We disentangle the different sources of complaining, finding that lack of communication and cash payment are the principal reasons of such lower ratings, with a negative incidence of respectively 27% and 21%.

We then analyzed the ratings posted along all the reviews and conduct a random-effects-logit regression to see whether the posted ratings are affected by the tax rates. We compare two scenarios in which costumers react either to the legal tax based on hotel stars or to the ad valorem tax on the room price. No effect is found on the legal tax rates, showing that what matters for customers' satisfaction is the weight of the tax on the price, i.e. the ad valorem tax. We estimate a 4% decrease in the probability of getting a review equal to 5 as the percentage tax rate goes from 1% to 10%. By allowing the effect to vary with the average price charged by hotels, we see that the negative effect is concentrated in cheaper hotels, where a standard room costs less than 100 Euros.

Our results are important both for firms operating in the tourist sector and for the public authority. Since the lack of information is the principal reason of the lower score for tax

mentioning reviews, hotel managers should be careful to advertise the existence of the tax and pursue transparent communication strategies. Current policies on tourist tax are based on calibrating legal tax rates in order to balance the need of financing local budgeting and the need of maintaining the competitiveness of the local destination. Across Europe, tourist tax rates are usually set on a per-person-per-night basis and are allowed to vary across hotel categories. We show that actually what matters for costumers is the weight of the tax on the price of the stays. As a consequence the actual tax system can create distortions as the hotels trying to keep a low price are penalized by an higher incidence of the tax. This suggests that local authorities should try to change the occupancy tax system and to introduce an ad valorem tax system, in which they are able to calibrate the tax rates considering their incidence on the price of the stay. In our sample, we show also that high-star costumers are almost insensitive to the level of the actual tax rates, whose incidence on price stays is lower that one-star hotels. Our results suggest that an increase in the percentage tax rate of high quality hotels should results in equity and efficiency gains for local jurisdictions, leading not only to increasing tax revenues but also to a more equal distribution of the tax burden.

References

Agrawal, D. R., Hoyt, W. H., and Wilson, J. D. (2021). Local Policy Choice: Theory and Empirics. *Journal of Economic Literature*.

Aguiló, E., Riera, A., and Rosselló, J. (2005). The short-term price effect of a tourist tax through a dynamic demand model. The case of the Balearic Islands. *Tourism Management*, 26(3):359–365.

Besley, T. and Case, A. (1995). Does Electoral Accountability affect Economic Policy Choices? Evidence from Gubernatorial Term Limits. *The Quarterly Journal of Economics*, (August).

Biagi, B., Brandano, M. G., and Pulina, M. (2017). Tourism taxation: A synthetic control method for policy evaluation. *International Journal of Tourism Research*, 19(5):505–514.

Bonham, C. S. and Gangnes, B. (1996). Intervention analysis with cointegrated time series: the case of the Hawaii hotel room tax. *Applied Economics*, 28:1281–1293.

Borges, A. P., Vieira, E., and Gomes, S. (2020). The Evaluation of Municipal Tourist Tax Awareness: the case of the city of Porto. *Tourism and Hospitality Management*, 26(2):381–398.

Candela, G., Castellani, M., and Mussoni, M. (2013). Imposta di soggiorno : spesa pubblica e tassazione nelle destinazioni turistiche. *Politica economica*.

Chen, C., Wu, S., and Xie, K. L. (2016). Online Consumer Review Factors Affecting Offline Hotel Popularity: Evidence from Tripadvisor. *Journal of Travel & Tourism Marketing*, 33(2):211–223.

Chevalier, J. A. and Mayzlin, D. (2006). The effect of word of mouth on sales: Online book reviews. *Journal of Marketing Research*, 43(3):345–354.

Chiarello, A. (2013). Le imposte comunali sul turismo. Technical report, IFEL Fondazione ANCI.

Denicolò, V. and Matteuzzi, M. (2000). Specific and Ad Valorem Taxation in Asymmetric Cournot Oligopolies. *International Tax and Public Finance*, 7:335–342.

Do Valle, P. O., Pintassilgo, P., Matias, A., and André, F. (2012). Tourist attitudes towards an accommodation tax earmarked for environmental protection: A survey in the Algarve. *Tourism Management*, 33(6):1408–1416.

Duràn-Romàn, J. L., Càrdenas-Garcìa, P. J., and Pulido-Fernandez, J. I. (2021). Tourist Tax to Improve Sustainability and the Experience in Mass Tourism Destinations : The Case of Andalusia (Spain). *Sustainability*, pages 1–20.

European Commission (2017). The Impact of Taxes on the Competitiveness of European Tourism. Technical report, European Commission.

Gao, B., Li, X., Liu, S., and Fang, D. (2018). How power distance affects online hotel ratings: The positive moderating roles of hotel chain and reviewers' travel experience. *Tourism Management*, 65:176–186.

Gooroochurn, N. and Sinclair, M. T. (2005). Economics of tourism taxation: Evidence from Mauritius. *Annals of Tourism Research*, 32(2):478–498.

Ho, R. C., Withanage, M. S., and Khong, K. W. (2020). Sentiment drivers of hotel customers: a hybrid approach using unstructured data from online reviews. *Asia-Pacific Journal of Business Administration*.

Rodella, I., Madau, F., Mazzanti, M., Corbau, C., Carboni, D., Utizi, K., Simeoni, U., and Saragat, V. (2019). Willingness to pay for management and preservation of natural , semi-urban and urban beaches in Italy. *Ocean and Coastal Management*, 172(January):93–104.

Schubert, S. F. (2010). Coping with externalities in tourism : a dynamic optimal taxation approach. *Tourism Economics*, 16(2):321–343.

Vermeulen, I. E. and Seegers, D. (2009). Tried and tested: The impact of online hotel reviews on consumer consideration. *Tourism Management*, 30(1):123–127.

Wang, K.-c. A., Chou, P.-y., and Liang, W.-j. (2018). Specific versus ad valorem taxes in the presence of cost and quality differences. *International Tax and Public Finance*, 25(5):1197–1214.

Wolf, C., Schmidt, P., and Stenzel, A. (2020). Pricing for the Stars. *Working Paper*.

Xiang, Z. and Gretzel, U. (2010). Role of social media in online travel information search. *Tourism Management*, 31(2):179–188.

A Ordered logit regressions

In this section we repeat the analysis made in Section 4 by using a random-effects-ordered-logit model. This methodology allows us to take into account the full discrete distribution of our dependent variable. In fact online ratings range from 1 to 5, but it is not reasonable to assume that the distance, in terms of tourist's satisfaction, from a rating of 1 to a rating of 2 equals the distance from 4 to 5. Ordered logit allows to account for this possible asymmetry in the measurement scale embedded in online ratings. By using a random effects model we account for possible unobserved hotel-level components which can influence our dependent variable.

The ordered-logit model can be represented as a latent-variable model. Let us call $y_{h,i}$ the ith rating posted about hotel h, with $y_{h,i} \in \{1, 2, 3, 4, 5\}$. Let us define $y^*_{h,i}$ as the continuous latent variable which maps to $y_{h,i}$ according to:

$$y_{h,i} = \begin{cases} 1 & \text{if} \quad y^*_{h,i} \leq \kappa_1 \\ 2 & \text{if} \quad \kappa_1 < y^*_{h,i} \leq \kappa_2 \\ \vdots & \\ 5 & \text{if} \quad \kappa_4 < y^*_{h,i} \end{cases} \tag{4}$$

where $\kappa_1, \kappa_2, ...$ are parameters to be estimated. The latent variable $y^*_{h,i}$ is assumed to be a linear function of all the covariates:

$$y^*_{h,i} = \alpha + \beta TaxRate_{h,i} + \pi X_{i,h} + u_h + \varepsilon_{i,t} \tag{5}$$

$$\pi X_{i,h} = \psi Price_h + \sum_j \gamma_j DType_h + \sum_k \delta_k DYear_{h,i} + \sum_m \tau_m DMonth_{i,h} + \sum_l \phi_l DCity_h \tag{6}$$

Figure 12 displays how predicted probabilities associated to each value of the rating varies with the tax rates, which in the graph are allowed to range from 1% to 5%. Vertical bars show 90% confidence intervals. We estimate a negative correlation between the probability that ratings take value 5 and the tax rate. The correlation is instead positive for all the other possible values of the ratings (from 1 to 4). This result confirms the evidence of a negative effect of higher tax rates on ratings, if we consider all ratings below 5 as indicating a negative outcome, though of different magnitude.

Figure 13 shows the same results when the official tax amount is used as regressor in place of the computed tax rate. No significant effect is found is this case.

Figure 14 shows how the estimated probabilities of ratings being equal to 5 are affected by the tax rate across the different hotel categories. These results are obtained by running

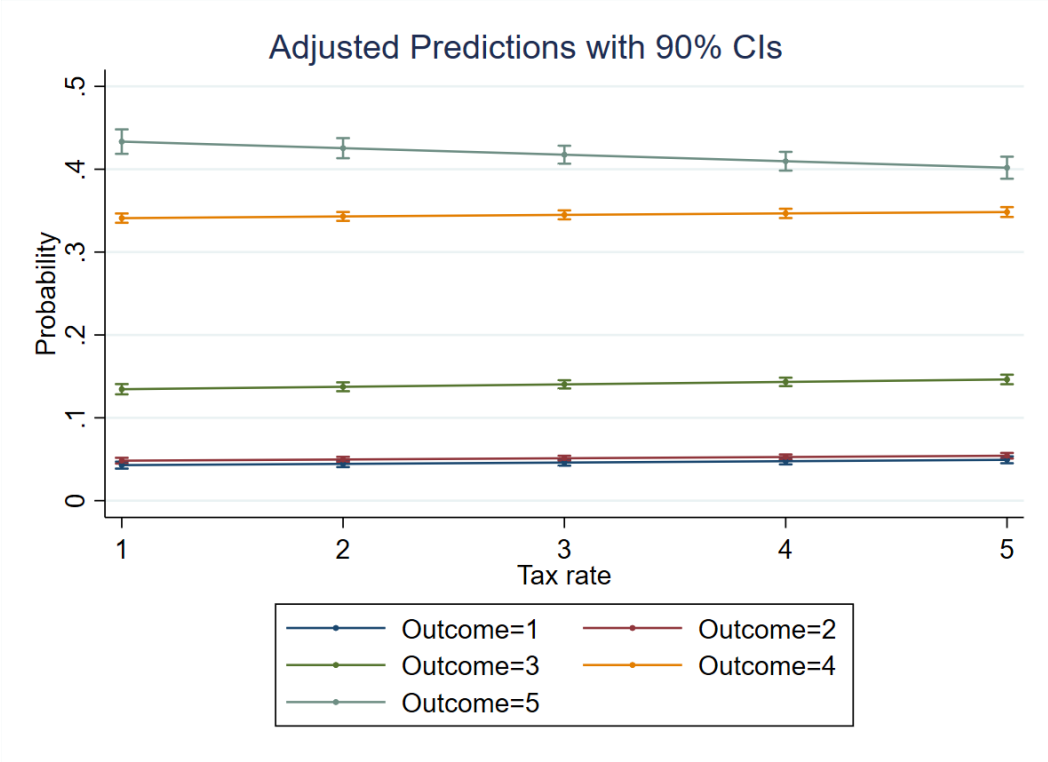

Figure 12: Estimated probabilities for each value of ratings resulting from the logistic regression versus the tax rate. Vertical bars show 90% confidence intervals.

separate logit regressions for each hotel category. The Figure shows that the negative effect is large significant for 3-stars hotels, while it completely disappears for 5-star hotels. For 3-star hotels, which constitute the large majority of hotels in Italy, we estimate a 5% reduction in the probability of getting an online rating equal to 5 as the tax rate goes from 1% to 5%.

We then let the β coefficient in Equation (5) to vary with the price. Figure 15 shows the results. We plot the difference between the estimated probability of having a rating equal to 5 when the tax rate is 5% versus the probability of having a rating equal to 5 when the tax rate is only 1%. This difference is shown for different levels of the price of a room. The estimated pattern is very clear: as the price of a room increases the negative effect on online ratings gets weaker. For prices of 100 Euros or higher, the effect is not significant anymore, confirming the results found above.

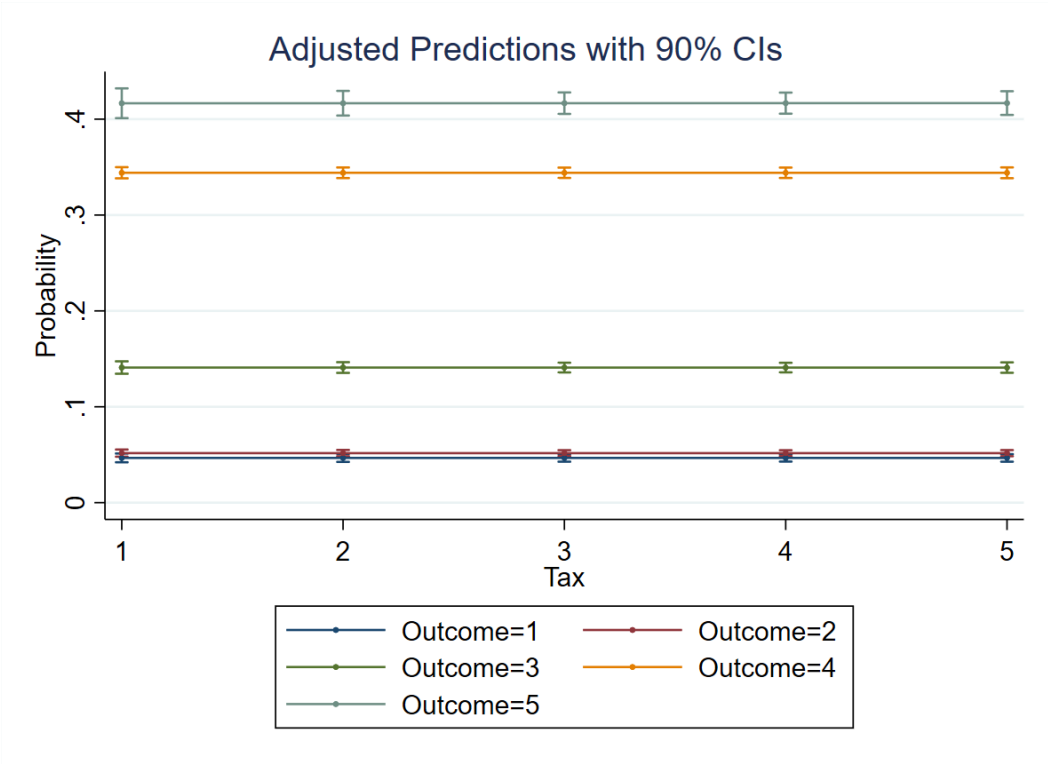

Figure 13: Estimated probabilities for each value of ratings resulting from the logistic regression versus the tax amount in Euros (per person per night). Vertical bars show 90% confidence intervals.

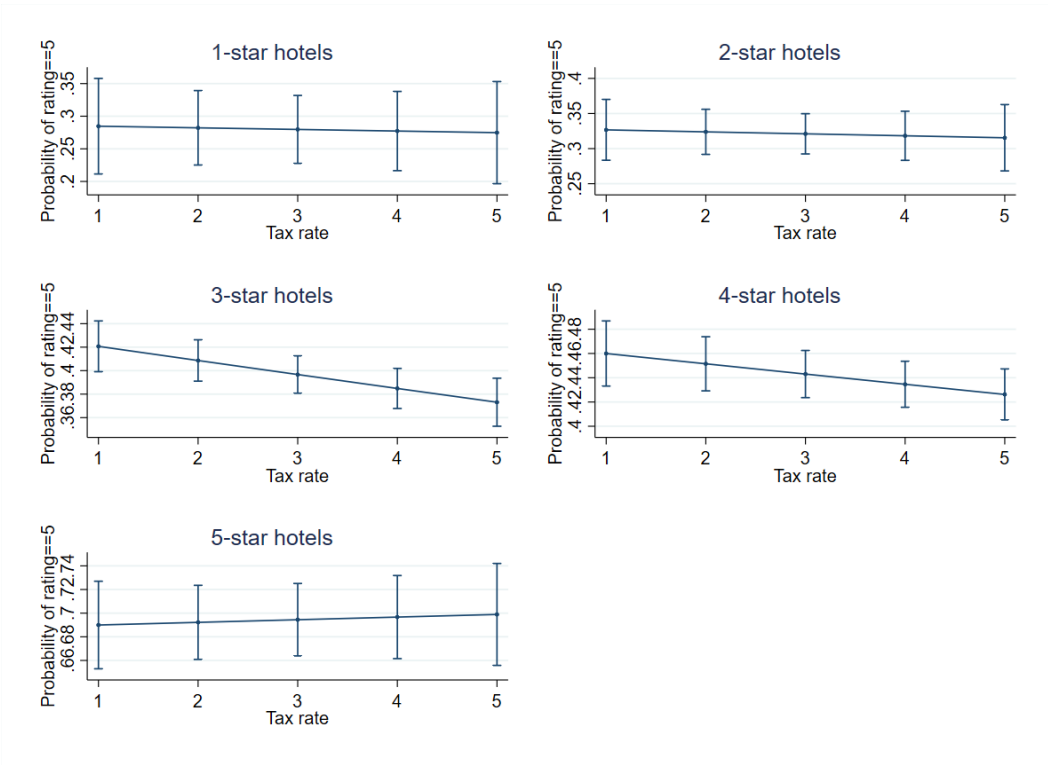

Figure 14: Estimated probabilities of rating being equal to 5 versus the tax rate, across hotel categories.

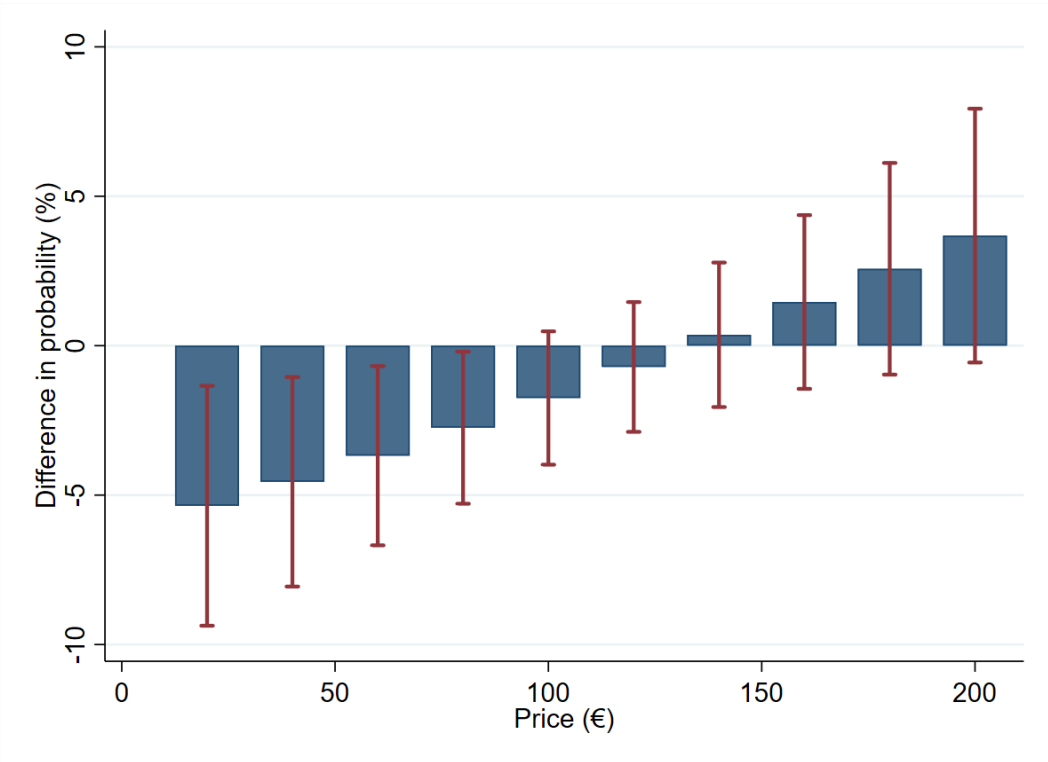

Figure 15: On the y-axis the estimated difference in probability $\hat{P}(Rating = 5/Tax\ rate = 5\%) - \hat{P}(Rating = 5/Tax\ rate = 1\%)$ is depicted, for different values of the price of a standard room. Vertical red bars denote 90% confidence intervals.

B Full results

Table 4 shows the full results from the estimation of the random effects logistic model explained in Section 5 (see equations from (1) to (3)). The Table reports the estimated odd ratios. An odd ratio greater (smaller) than one implies a positive (negative) effect of the regressor on the probability of a positive outcome, i.e. that the rating is equal to 5.

Models (1) and (2) differ only in the relevant variable used for measuring the tax burden, which is the legal tax rate (on a per person per night basis) in model (1) and the % tax rate in model (2). The coefficients related to these two variables are extensively discussed in Section 5. In this Appendix we take a look also at the estimated coefficients on the other control variables. Regarding the hotel categories, a positive and significant coefficient is estimated only for 4-stars and 5-stars hotels (1-star hotel is the reference category). Nonetheless, we can see that: (i) the estimated odd ratios are all greater than 1, which means that all categories display a better conditional rating compared to the 1-star category; (ii) the estimated odd ratios are increasing with the number of stars, suggesting a monotonic relationship between the number of stars and the conditional probability of receiving a rating equal to 5.

By looking at the estimated odd ratios on the dummies referring to the year in which the reviewer has visited the hotel, we observe an increasing trend over the years: more recent reviews display an higher conditional probability of having a rating equal to 5.

Among the cities in the sample, only for Milano and Rimini we observe a significant difference with respect to the reference (Firenze), which display respectively a lower and higher conditional probability.

As expected, the price variable, being very much correlated with the unobserved quality of the hotel, has a positive impact on the estimated probability of having a rating equal to 5.

The coefficients associated to the month dummies have not been included in the Table for space purposes, but the estimated seasonal effects are not significant for almost all months, except for May and October, for which we estimate a negative impact, and for August, for which we observe a positive impact.

	(1)	(2)
Legal tax rate	0.9998	
	(0.9890)	
% tax rate		0.9802**
		(0.0064)
Type=2-stars hotel	1.0422	1.0537
	(0.8262)	(0.7752)
Type=3-stars hotel	1.2654	1.2967
	(0.1776)	(0.1237)
Type=4-stars hotel	1.4403*	1.5140*
	(0.0480)	(0.0190)
Type=5-stars hotel	2.2800**	2.4607***
	(0.0027)	(0.0007)
Year=2012	1.0687	1.0844
	(0.1868)	(0.1062)
Year=2013	1.0934	1.1262*
	(0.0920)	(0.0238)
Year=2014	1.1817**	1.2508***
	(0.0036)	(0.0001)
Year=2015	1.1374*	1.2335***
	(0.0399)	(0.0006)
Year=2016	1.2217**	1.3262***
	(0.0023)	(0.0000)
Year=2017	1.3170***	1.4331***
	(0.0000)	(0.0000)
Year=2018	1.3555***	1.4764***
	(0.0000)	(0.0000)
Year=2019	1.4596***	1.6003***
	(0.0000)	(0.0000)
City=Milano	0.5312***	0.5372***
	(0.0000)	(0.0000)
City=Napoli	1.1047	1.0666
	(0.5543)	(0.6973)
City=Rimini	1.8338***	1.7174***
	(0.0000)	(0.0000)
City=Roma	1.1689	1.1919
	(0.1768)	(0.1234)
Avg. price	1.0040***	1.0036***
	(0.0000)	(0.0000)
NT	126476	126476

Odds ratio p-values in parentheses

$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

Table 4: Estimated odd ratios. Random effects logit regressions.

Alma Mater Studiorum - Università di Bologna
DEPARTMENT OF ECONOMICS

Strada Maggiore 45
40125 Bologna - Italy
Tel. +39 051 2092604
Fax +39 051 2092664
http://www.dse.unibo.it