CESifo WORKING PAPERS

7972
2019
November 2019

How Does Competition Affect Reputation Concerns? Theory and Evidence from Airbnb

Michelangelo Rossi

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
www.cesifo-group.org/wp

An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: www.CESifo-group.org/wp

CESifo *Working Paper No. 7972*
Category 14: Economics of Digitization

How Does Competition Affect Reputation Concerns? Theory and Evidence from Airbnb

Abstract

I show how changes in competition affect the power of reputation to induce sellers to exert effort. The impact of competition on sellers' incentives is theoretically ambiguous. More competition disciplines sellers, but, at the same time, it erodes reputational premia. This paper identifies empirically whether one effect dominates the other using data from Airbnb. To guide the empirical analysis, I develop a model of reputation where the relative number of hosts and guests affects the value of building a reputation through effort. In this framework, more competition depresses hosts' profits and leads hosts to reduce effort. I test the model's prediction exploiting a change in regulation for Airbnb listings effective in San Francisco in 2017. I identify a negative causal effect of competition on effort. As the number of competitors surrounding each listing increases by 10 percent, ratings about hosts' effort decrease by more than one standard deviation. These findings suggest that more competition may erode incentives for high-quality services in markets where sellers' performances depend on reputation.

JEL-Codes: D820, D830, L500, L810.

Keywords: reputation, competition, digitization.

Michelangelo Rossi
Universidad Carlos III
Department of Economics
Madrid / Spain
mrossi@eco.uc3m.es

Click here:
https://sites.google.com/view/michelangelorossi/research?authuser=0 for the latest version.

November 9, 2019
I am grateful to my advisors, Natalia Fabra and Matilde Pinto Machado, for their patient guidance and advice. I also thank Victor Aguirregabiria, Heski Bar-Isaac, Jesús Carro, Ambarish Chandra, Rahul Deb, Alexandre de Cornière, Daniel Ershov, Chiara Farronato, Alberto Galasso, Ignatius J. Horstmann, Bruno Jullien, Gerard Llobet, Luigi Minale, Matthew Mitchell, Maryam Saeedi, Jan Stuhler, Mark Tremblay, and Gabor Virag. Helpful feedback was received at the internal seminars at Universidad Carlos III de Madrid, CEMFI, and Rotman School of Management, at the International Industrial Organization Conference, the Doctoral Workshop on The Economics of Digitization, the CEPR/JIE School on Applied Industrial Organisation, the European Association Research Industrial Economics Conference, the Jornadas de Economia Industrial, and the TSE Digital Workshop. Special gratitude to Elizaveta Pronkina for her enormous support and encouragement. All errors are mine. Support from the Ministerio Economía, Industria y Competitividad (Spain), grant ECO2016/78632-P, is gratefully acknowledged.

1 Introduction

The rise of online marketplaces had a strong impact on many industries in the last decade. Thanks to their low entry costs, these platforms facilitate the entry of new service providers. As an example, in the last years Airbnb hosts massively entered the hospitality market and now they outnumber the largest hotel chains in terms of available rooms: in 2019, there are more than seven million Airbnb listings around the world. For comparison, the Marriott chain has less than 1.5 million rooms.[1] Digital platforms rely on review systems to ensure the quality of their services and to provide incentives for sellers to exert effort. Reviews reveal sellers' past performances and form their reputation. Thus, sellers' concerns for a good reputation are one of the key ingredients for the quality of online transactions and the success of digital marketplaces.

This paper studies how changes in the number of competitors affect the power of reputation to provide incentives for sellers to exert effort. The effects of competition on the sellers' incentives are theoretically ambiguous (Bar-Isaac, 2005). More competition may help to discipline sellers, but, at the same time, it erodes reputational premia. Understanding which of the two effects dominates empirically is a relevant question, not only for the design of digital platforms, but also for other markets in which sellers' quality is unknown and their performances depend on reputation. This feature is common to several markets involving experience goods and services such as hospitals, restaurants, or schools. Yet, the process of reputation building is particularly relevant in online marketplaces since review systems provide an observable measure of sellers' reputation and effort.

I empirically address this research question using data from one of the fastest-growing online platforms: Airbnb. This setting is of special interest since the enormous growth in the number of hosts on the platform has attracted considerable attention from local governments and regulators. Previous works have shown that the entry of Airbnb hosts in a city expands the number of available rooms, reduces hotels' profits, and increases consumers' welfare (Zervas et al., 2017; and Farronato and Fradkin, 2018). Yet, in addition to the disciplining impact of competition on prices, an increase in the number of competitors may also impact hosts' incentives to exert effort affecting the quality of platform's services. To the best of my knowledge, this is the first paper to identify the causal impact of the number of competitors on the incentives to exert effort in online marketplaces: my findings show that, when the number of competitors increases, hosts exert less effort and their profits reduce.

To inform my empirical analysis, I develop a model of reputation in which the number of hosts and guests on the two sides of the market (the market tightness) impacts the reputation return of hosts' effort. The model predicts that, when the number of competitors decreases, hosts'

[1] For more information about the recent growth of Airbnb around the world, see the annual official reports provided by Airbnb at https://press.airbnb.com/fast-facts/.

profits increase and hosts exert more effort. With fewer competitors, hosts have higher probability to be matched with a guest and can charge higher prices. However, the price elasticity of hosts' demand depends on their reputation. In particular, the probability to be matched with a guest is less elastic for hosts with good reputation. Accordingly, the premium of exerting effort (and having good reputation) increases when the number of competitors is lower: hosts with good reputation can post higher prices with a lower reduction in their demand.

In order to test the model's prediction, I analyze empirically the relationship between the effort exerted by Airbnb hosts and the number of their competitors. I measure hosts' effort by studying ratings such as *communication* and *check-in* that are specifically related to hosts' actions. Moreover, to measure the number of competitors for each host, I create host-specific consideration sets by counting the number of listings surrounding each host within a radius of 0.5, 1, and 2 kilometers. Doing so, I assume that Airbnb hosts compete more strongly with listings that are closely located to them, relative to those further away. This is in line with Zervas et al. (2017) who show that the impact of Airbnb entry on hotels' revenues is sensitive to the distance between hotels and Airbnb listings.

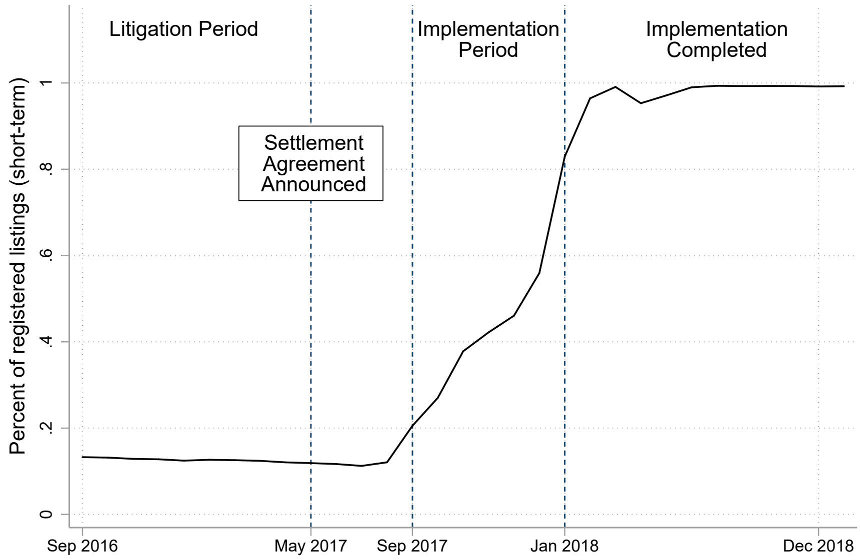
My identification strategy exploits a unique quasi-experiment to isolate the effect of changes in the number of competitors from other confounders. In particular, I take advantage of a regulatory enforcement on short-term rentals that occurred in San Francisco in 2017.

Airbnb was founded in San Francisco. In 2015, it was the third US city in terms of active Airbnb listings after New York and Los Angeles (Lane et al., 2016). From that year, the San Francisco City Council imposed several restrictions and a formal registration for short-term rentals on digital platforms.[2] Yet, the regulation started to be effectively enforced only two years later, when Airbnb signed a Settlement Agreement with the City Council in May 2017. Accordingly, the platform has been actively engaged in the listings' registration process since September 2017. As shown in Figure 1, the percentage of Airbnb listings offering short-term lodging formally registered at the City Council Office dramatically increased from less than 15 percent in September 2017 to 100 percent in February 2018: hosts started to register, and those who could not, exited the platform. As a result, a few months after September 2017, the number of Airbnb listings offering short-term lodging halved, dropping from about 8,000 units in September 2017 to less than 4,000 in February 2018 (see Figure 2).

I exploit this regulatory enforcement as an exogenous shift in the number of listings surrounding each host. I focus on hosts renting short-term that were present on the platform both before and after the Settlement Agreement. By such selection, I abstract from hosts' decision to enter or exit due to the regulation enforcement. All hosts renting short-term in San Francisco are affected

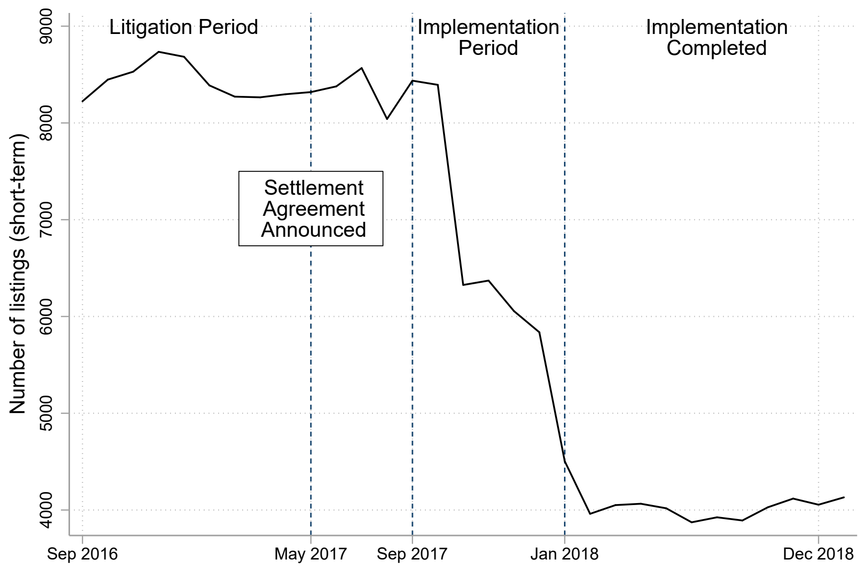
[2]Rentals are considered "short-term" if the properties are rented for less than 30 consecutive nights at a time.

Figure 1: Percent of Registered Airbnb Listings over Time



Note: The figure plots the percentage of Airbnb listings offering short-term lodging that displayed a registration number in San Francisco over time from September 2016 to January 2019. The three vertical lines regard the timing of the Settlement Agreement between the San Francisco City Council and Airbnb. The agreement was signed and announced in May 2017 and it has been effective since September 2017. According to the resolution, from January 2018 all eligible Airbnb listings should be registered.

Figure 2: Short-term Airbnb Listings over Time



Note: The figure plots the total number of Airbnb listings offering short-term lodging in San Francisco over time from September 2016 to January 2019. The three vertical lines regard the timing of the Settlement Agreement between the San Francisco City Council and Airbnb. The agreement was signed in May 2017 and it has been effective since September 2017. According to the resolution, from January 2018 all eligible Airbnb listings should be registered.

by the Settlement Agreement. On the other hand, the exposure to this "shock" differs since the variation in the number of competitors is heterogeneous across hosts. I take advantage of this heterogeneity in the treatment. To measure the exposure of each host, I use the percentage of listings surrounding each host that were already registered in September 2017. For higher values of this percentage, fewer listings are likely to exit after the Settlement Agreement since they were already complying with the regulation. I employ this measure as a predictor for the variation in the number of listings surrounding each host after the Settlement Agreement. Therefore, I identify the effect of variations in the number of competitors by using the differential changes in the exposure across listings over time. The core identifying assumption of this design shares the intuition of a difference in differences estimator with a continuous treatment. However, the identification s based on an instrumental variable regression where the excluded instrument is given by the interaction between the measure of exposure (the percentage of listings already registered in September 2017) and time.

The results show a statistically and economically significant negative relationship between the number of competitors and hosts' effort. When the number of competitors decreases by 10%, ratings regarding hosts' effort increase by more than one standard deviation. I corroborate this result studying variations in hosts' profits and in the monetary value of reputation. With the same identification strategy, I find that less competition increases profits, so that hosts have higher incentives to exert effort. Moreover, since fewer hosts are going to have good reputation, I show evidence that the signaling effect of reputation is stronger in more competitive frameworks.

This paper makes contributions to both the empirical and theoretical literatures on reputation. To the best of my knowledge, there are no studies that investigate empirically the relationship between competition and sellers' reputational incentives to exert effort. The empirical literature about reputation has grown in past years with a particular focus on online settings. Still, the majority of the empirical papers study the impact of online feedback on sellers' profits and they do not specify the mechanism behind this effect. Cabral (2012) and Tadelis (2016) give excellent and comprehensive reviews of the empirical literature on this topic.

The closer paper to mine is Elfenbein et al. (2015). They study the effect of quality certification on the probability to sell an item for eBay sellers. Their results show that the positive effect of certification is higher in more competitive settings and when certification is scarce. They do not specifically study sellers' incentives to exert effort, although their main result is in line with a negative relationship between the number of competitors and sellers' effort. With more competition, fewer sellers exert effort, thus good reputation is more scarse and its signaling power is higher. The online setting I use to address my question (Airbnb) presents clear methodological advantages relative to Elfenbein et al. (2015). Thanks to the multiple components of the Airbnb review system, I can use ratings regarding *communication* and *check-in* as a proxy for hosts' effort. Moreover, thanks to the information regarding the geographical location of each Airbnb host, I can exploit the heterogeneous impact of the regulation enforcement for causal identification.

Only recently, a few papers analyze the role of review systems to reduce asymmetries of information in contexts with adverse selection, or moral hazard. These papers focus on the design of online review systems and they do not study the role of competition on the sellers' incentives to exert effort. Klein et al. (2016) and Hui et al. (2016) take advantage of a variation in the eBay review system implemented in 2008 to study changes in eBay sellers' performance. The modification in the review system reduced buyers' fear of retaliation by sellers and improved the transparency of the online feedback. While Klein et al. (2016) claim that this change induced a disciplining effect on sellers' behavior (moral hazard), Hui et al. (2016) attribute the improvement to seller's selection (adverse selection). In the Airbnb setting, Proserpio et al. (2018) show that members' reciprocity is relevant and users can induce others to behave well by exerting more effort themselves.

From a theoretical perspective, the relationship between competition and sellers' incentives to exert effort is ambiguous. To guide my empirical strategy, I present a model of reputation building that embodies some of the most salient characteristics of digital platforms: search frictions. The model highlights the theoretical mechanism behind the empirical results and it helps to connect the variations in the number of competitors with hosts' effort, profits, and the value of reputation. A few theoretical papers have investigated the relationship between competition and sellers' incentives to exert effort. Most previous reputation models studied the repeated effort choices by a long-lived monopoly seller meeting short-lived buyers in every period. For a comprehensive review of the theoretical literature regarding reputation, see Bar-Isaac and Tadelis (2008). I am aware of only two papers, Kranton (2003) and Bar-Isaac (2005), that explicitly investigate how variations in the extent of competition affect sellers' incentives to exert effort. Kranton (2003) studies the decisions to provide high or low quality goods by a finite number of firms competing in a repeated game. She assumes that, after a firm produces low quality, its future profits are null, independently of competition. Accordingly, an increase in the number of competitors only reduces the benefits of having reputation for high quality and it results in lower incentives to exert effort. Bar-Isaac (2005) allows firms' profits to depend on the number of competitors after a firm produces low quality. As a result, the effects of competition on effort are ambiguous. With a higher degree of competition, profits with reputation for low quality are lower (competition disciplines agents), but, at the same time, profits with reputation for high quality are also lower (competition erodes reputational premia). In contrast with these two papers, my model considers a directed search framework where the matching between hosts and guests is frictional. This is in line with the recent empirical research on online marketplaces that emphasizes how digital platforms are inherently frictional settings (Fradkin, 2015, Fradkin, 2017, and Horton, 2019). Guests direct their search to hosts after observing prices and their past effort choices. Accordingly, hosts who exerted effort in the past can charge higher prices and have higher probability to be matched with guests. Conversely, hosts who did not exert effort have to charge lower prices to make guests indifferent at the moment of choosing where to direct their search. However, the price elasticity of hosts' probability to be matched with a guest depends on hosts' reputation. In particular, using standard assumptions on the matching function between hosts and guests, I can show that hosts' matching probability is less elastic to price changes when hosts have good reputation: hosts can post higher prices suffering a lower reduction in their demand when they have good reputation. Therefore, in my model, more competitors lead to lower profits independently of the current hosts' reputation (as in Bar-Isaac (2005)). Yet, the negative effect is stronger for hosts who did not exert any effort in the past: hence, competition erodes the power of reputation to discipline hosts' behavior.

Outside the literature on reputation, several papers analyze the effects of competition on firms' investment decision. Aghion et al. (2001) and Aghion et al. (2005) study the relationship between product market competition and innovation and show empirical evidence of an inverted-U relationship using aggregate data on several industries. In monopolistic settings, firms' investments are low and more competition is beneficial for innovation. Yet, when the starting level of competition is high, an increase in competition may be detrimental. From this perspective, my paper studies a specific type of investment (hosts' effort) that each host decides to make at every transaction, and whose returns are only in terms of reputation. Accordingly, the contribution of this paper to the literature regarding competition and investment is twofold: first, I analyze a context (digital platforms) and a type of investment (hosts' effort) that have never been studied before. Then, I provide a methodological contribution since I identify the causal impact of the number of competitors on hosts' effort exploiting a unique quasi-experiment.

The rest of the paper is organized as follows. Section 2 describes the theoretical model and the testable predictions. In Section 3, first I provide some background context regarding Airbnb. Then, I illustrate the change in the institutional setting regarding Airbnb hosts regulation in San Francisco in September 2017. Next, I present the dataset. I discuss my identification strategy in Section 4. Section 5 provides the main empirical findings. In Section 6, I show further results in line with the theory. I proceed with the robustness checks in Section 7. Section 8 concludes. All the proofs and additional tables are in Appendix.

2 Model

In this Section, I present the theoretical framework underlying my analysis. First, I describe the model environment. I show the agents' characteristics and payoffs; and I clarify the role of frictions with the assumptions regarding the matching function. Then, I present the timing of agents' interactions and the equilibrium concept. Finally, I characterize the equilibrium allocation and propose the main testable predictions of the model. All proofs are in Appendix A.

2.1 Model Environment

The market lasts two periods. Hosts and guests populate the two sides of the market. Each guest (he) is willing to rent a house, whereas each host (she) owns a house and can rent it to one guest only. In both periods there is an infinite population of hosts who can potentially enter the market. Hosts who enter in the first period stay in the market until the second period. To enter the market, hosts pay entry costs, f, in both periods. Once she entered, a host posts a price p, and, in case of a match with a guest, decides whether to exert effort or not: $e = \{0, 1\}$. A host's cost of effort, c, is realized if a host is matched and it is permanent across periods. The cost can take two values: $c = \{0, k\}$ with $k > 0$. Hosts draw $c = 0$ with probability π. The cost is the host's private information, whereas the probability π is common knowledge for hosts and guests. A unit mass of guests is present in the market in period 1; instead, a measure G is present in period 2. Guests are homogeneous and the gross utility from a transaction, u, depends on host effort and price: $u = ae + b - p$, with $a, b \geq 0$. b represents the benchmark utility that guests obtain from a transaction when hosts do not exert effort. The ex-post surplus of a transaction is defined by the sum of guest's utility and hosts' profit. If the host exerts effort, $e = 1$, the ex-post surplus is $(a + b - p) + (p - c) = a + b - c$. If the host does not exert effort, $e = 0$, the ex-post surplus is $(b - p) + p = b$. In order to guarantee the efficiency of exerting effort $e = 1$, I assume that $a - c > 0$ and that hosts always exert effort $e = 1$ if they draw $c = 0$.

The matching process between hosts and guests is frictional. In line with the directed competitive search literature, market frictions are captured by a matching function M. With a measure h of hosts and g of guests present in the market, a measure $M(h, g) \leq \min(h, g)$ of matches is formed. Assuming constant returns to scale in the matching function, the agents' probability of transacting can be determined as a function of the ratio between guests and hosts, denoted as the market tightness: $\theta = \frac{g}{h}$.

The hosts' probability of transacting when the market tightness is θ is defined as $\alpha(\theta) \equiv \frac{M(h,g)}{h}$. Whereas the guests' probability is defined as $\frac{\alpha(\theta)}{\theta} \equiv \frac{M(h,g)}{g}$. I impose the following conditions on the function $\alpha(\theta)$:

Assumption 1. *For all $\theta \in [0, \infty)$:*

1. *$\alpha(\theta) \in [0, 1]$ and $\frac{\alpha(\theta)}{\theta} \in [0, 1]$;*

2. *$\alpha(\theta)$ is continuous, strictly increasing, twice differentiable, and strictly concave;*

3. $\alpha(\theta) - \theta\alpha'(\theta) > 0$;

4. $\alpha(\infty) = \alpha'(0) = 1$ *and* $\alpha(0) = \lim_{\theta\to\infty} \theta\alpha'(\theta) = 0$.

Assumption 1 is standard in the directed search literature[3]. In particular, $\alpha'(\theta) > 0$ and $\alpha(\theta) - \theta\alpha'(\theta) > 0$ state that, when the number of guests over hosts increases, the host matching probability strictly increases and the guest matching probability strictly decreases. The expected payoffs of hosts and guests can be defined in terms of the host effort and pricing decisions and the probability of having a transaction. In each period, the expected profit for hosts is:

$$\Pi = (p - ce)\alpha(\theta);$$

whereas the expected utility for guests is:

$$U = (ae + b - p)\frac{\alpha(\theta)}{\theta}.$$

The timing of the model is the following. In period 1:

1. Each host decides to enter the market;
2. Each host posts price: $p_1 \in \mathbb{R}^+$;
3. Guests form beliefs about the hosts' expected effort decision observing p_1: $\mu_1(p_1)$;
4. Guests choose where to direct their search and matches are formed;[4]
5. Each host matched with a guest draws the cost of effort c;
6. Each host chooses whether or not to exert effort: $e_1(c, p_1)$;
7. Transactions occur.

At the end of period 1, a history h is formed for each host and it is public information. If the host had a transaction, her history is composed by the effort exerted, $h = (e_1(c))$. If the host did not have a transaction, her history is composed by the information that the host had no guests: $h = (\emptyset)$. Hosts who enter in period 2 have a blank history $h = (\emptyset)$.

After observing histories, guests form interim beliefs $\bar{\mu}_2(h)$ about hosts effort decision in next period.

[3] Delacroix and Shi (2013) and Shi and Delacroix (2018) extensively discuss the class of matching functions satisfying Assumption 1 in the literature.

[4] I do not explicitly model the search process by guests. Depending on how the market is organized, different matching functions (all satisfying Assumption 1) can be micro-founded. For further details, see Peters (1991), Burdett et al. (1995) and Burdett et al. (2001).

In period 2, the same timing applies. However, guests update their interim beliefs about hosts' effort observing current prices:

1. Each host decides to enter the market;
2. Each host posts price: $p_2(c, h) \in \mathbb{R}^+$;
3. Guests update interim beliefs about hosts' expected effort decision observing h and $p_2(c, h)$: $\mu_2(p_2(c, h), h)$
4. Guests choose where to direct their search;
5. Each host matched with a guest who was not matched in period 1 draws the cost of effort c;
6. Each host chooses whether or not to exert effort: $e_2(c, p_2(h), h)$;
7. Transactions occur.

2.2 Equilibrium Characterization

The equilibrium concept used is symmetric perfect Bayesian equilibrium with pure strategies in prices. In this setting, posted prices play two separate functions. First, prices "direct" guests' search behavior as they affect the number of guests who are willing to be matched with hosts. Moreover, prices posted in period 2 can be a signal for hosts' cost of effort. I limit my analysis imposing some assumptions regarding these two tasks of prices.

In line with the directed search literature, I assume that, in each period, the ex-ante guests' utility U_t cannot be affected by the price posted by a single host:

$$U_t = (a\bar{\mu}_t + b - p_t)\frac{\alpha(\theta_t)}{\theta_t}, \tag{2.1}$$

where $\bar{\mu}_t$ defines guests' beliefs about hosts' effort choice. Accordingly, changes in price p_t that do not affect guests' beliefs $\bar{\mu}_t$ are fully compensated by changes in tightness θ_t: if a host chooses a lower price, more guests direct their search towards her until the tightness increases and the guests' probability of transacting decreases. Equation 2.1 characterizes guests' beliefs about tightness levels for every price, even for those prices that are not posted in equilibrium. This approach is known in the directed search literature as the "market utility" approach (Wright et al., 2017).

In this setting, prices in period 2 can also serve as a signal for hosts' cost of effort since they can affect guests' beliefs $\bar{\mu}_t$. After a host is matched with a guest in period 1, her cost of effort is realized and it is private information. Hosts' cost of effort is relevant for guests' utility: while hosts with cost $c = 0$ always exert effort, hosts with positive cost $c = k > 0$ strategically choose whether to exert effort or not.

Yet, prices in period 2 are not the only variable signaling hosts' cost of effort. Hosts' histories are observed by guests in period 2 and they may be informative about hosts' cost. When a host's history reports $e_1 = 0$, guests in period 2 know with certainty that she has positive cost of effort (hosts with zero cost always choose to exert effort) and she does not exert effort in period 2: $\bar{\mu}_2 = 0$. Differently, histories reporting $e_1 = 1$ can sustain positive guests' beliefs about hosts' effort in period 2 ($\bar{\mu}_2 \geq \pi$).

The signaling functions of prices and histories are related. If prices fully solve the asymmetry of information between hosts and guests, histories' signal of hosts' cost of effort is ineffective. In particular, if hosts with different cost of effort have separate pricing strategies in period 2, then guests perfectly infer hosts' costs and, in equilibrium, hosts with zero cost exert effort $e_1(0) = e_2(0) = 1$, whereas hosts with positive cost do not exert effort $e_1(k) = e_2(k) = 0$. I restrict my analysis over a class of equilibria where histories provide effective signals about hosts' costs, and I denote these equilibria as *reputational equilibria*.[5] I focus on reputational equilibria for two reasons. First, empirical evidence suggests that prices do not fully reveal users' private type. Histories (reviews) are important to reduce the asymmetry of information in digital platforms.[6] Moreover, outside the class of reputational equilibria, hosts who draw a positive cost of effort in period 1 do not exert effort in any of the two periods ($e_1(k) = e_2(k) = 0$). Differently, in reputational equilibria, hosts who draw a positive cost may exert effort in period 1 ($e_1(k) = 1$) in order to mimic hosts with $c = 0$ and get a price premium in period 2. Thus, since exerting effort is efficient ($a > c$), reputational equilibria are Pareto superior in terms of the ex-post surplus of transactions relative to other non-reputational equilibria.

Pooling strategies in prices for hosts with the same history in period 2 characterize the class of reputational equilibria. In period 1, all hosts post the same price since the cost of effort is drawn after matches are formed. Accordingly, guests in both periods cannot infer hosts' costs directly from prices in period 1. After transactions occur, hosts have different histories depending on the reported effort, which affect guests' beliefs $\bar{\mu}_2$ about hosts' effort choice in the future. In period 2, hosts with the same history post the same price. In particular, hosts who were not matched in period 1 and new entrants post the same price since their cost of effort is drawn after matches. The case is similar for hosts who were matched in period 1. By pooling in prices, hosts with $c = k > 0$ obtain a price premium in period 2 if they exert effort in period 1. It constitutes the reputational benefit (the "carrot") of having exerted effort. Conversely, if hosts with $c = k > 0$ do not exert effort, they cannot

[5] In Appendix A, I discuss non-reputational equilibria and I show that their existence and stability rely on further assumptions regarding model's parameters.

[6] Cabral and Hortaçsu (2010), Fan et al. (2016), and Jolivet et al. (2016) show evidence regarding the significant impact of reviews on sellers' profitability in several online marketplaces.

pool in period 2 and their cost is fully disclosed (the "stick"). Price pooling is vital to implement the "carrot-stick strategy" that characterizes reputational equilibria. Multiple prices can sustain these equilibria and a continuum of equilibria is present in this class. In the main text, I restrict my analysis to the price profile that implements the constrained efficient equilibrium allocation and maximizes hosts' profits. To do so, I consider guests' beliefs that disregard the additional signaling role of prices in period 2: for any posted price, guests in period 2 do not update their beliefs about hosts' cost of effort (formed observing the host's history). This restriction is not necessary since a wide range of guests' beliefs sustains the constrained efficient equilibrium allocation. Disregarding the signaling from prices in period 2 is justified by the following observation. Independently of their cost of effort, hosts with the same history in period 2 have the same profit function: hosts with $c = k > 0$ do not exert effort in period 2 and their expected profits are $p_2\alpha(\theta_2)$; similarly, hosts with $c = 0$ do exert effort (that is costless for them) and get $p_2\alpha(\theta_2)$ as well. Accordingly, the optimal pricing strategy is aligned for both hosts' types and guests may not update their beliefs after observing prices in period 2. Furthermore, thanks to the equality of the profit function in period 2 for hosts with different costs of effort, reputational pooling equilibria are *not* eliminated by refinements such as the intuitive criterion by Cho and Kreps (1987).

Before providing a formal definition of the equilibrium, I characterize hosts' decisions proceeding by backward induction.[7]

2.3 Hosts' Decisions: Period 2

The effort decision in period 2 is straightforward.

Lemma 1 (Effort Decision in Period 2)**.** *In equilibrium, hosts who are matched with a guest in period 2 exert effort if and only if they have zero cost of effort $c = 0$.*

Lemma 1 directly follows from the assumption that hosts with cost $c = 0$ always exert effort. Differently, hosts with cost $c = k > 0$ always exert $e_2(k) = 0$ since effort is costly for them and they cannot commit to exert positive effort since guests direct their search without knowing hosts' effort decision.

In period 2, hosts post prices to match with guests. Hosts with the same history who were matched with guests in period 1 post the same price. Hosts who were not matched with guests in period 1 post the same price as new entrants since no information pertaining their cost of effort is revealed.

[7] In Appendix A, I illustrate the constrained efficient allocation and I discuss the Hosios (1990) conditions that characterize the equilibrium (proposed in the main text) implementing this allocation.

In the remaining part of the analysis, I use the following notation. I denote histories that appear in equilibrium with positive probability as follows:

$$\begin{aligned} h^1 &= (e_1 = 1); \\ h^0 &= (e_1 = 0); \\ h^{\emptyset} &= (\emptyset). \end{aligned}$$

Superscripts denote hosts' costs. I use superscript "pool" if hosts who draw different cost of effort may play the same strategy. If hosts who have not yet drawn the cost of effort play a strategy, I use superscript "$\emptyset$". Accordingly, the same notation $h^{\emptyset}$ can be used to denote histories for hosts who enter in period 1 and are not matched with guests; and for hosts who enter in period 2.

Proposition 1 (Pooling in Prices in Period 2)**.** *In any reputational equilibrium, hosts who were matched with a guest in period 1 and have the same history $h = \{h^0, h^1\}$ post the same price in period 2. Given guests' interim beliefs $\bar{\mu}_2$ and the expected utility U_2, hosts post prices $p_2^{pool}(h)$ and guests direct their search so as to form tightness $\theta_2^{pool}(h)$:*

$$\begin{aligned} \alpha'(\theta_2^{pool}(h)) &= \frac{U_2}{a\bar{\mu}_2(h) + b} \\ p_2^{pool}(h) &= a\bar{\mu}_2(h) + b - \frac{\theta_2^{pool}(h)}{\alpha(\theta_2^{pool}(h))} U_2, \end{aligned}$$

if $a\bar{\mu}_2(h) + b \geq U_2$. Otherwise, $\theta_2^{pool}(h) = 0$ and $p_2^{pool}(h) = 0$. Hosts who were not matched with a guest in period 1 and new entrants post the same price $p_2^{\emptyset}$ and guests direct their search so as to form tightness $\theta_2^{\emptyset}$:

$$\begin{aligned} \alpha'(\theta_2^{\emptyset}) &= \frac{U_2}{a\bar{\mu}_2(h^{\emptyset}) + b} \\ p_2^{\emptyset} &= a\bar{\mu}_2(h^{\emptyset}) + b - \frac{\theta_2^{\emptyset}}{\alpha(\theta_2^{\emptyset})} U_2, \end{aligned}$$

if $b \geq U_2$. Otherwise, $\theta_2^{\emptyset} = 0$ and $p_2^{\emptyset} = 0$.

The proof of this proposition is in Appendix A. Proposition 1 establishes a relationship between the price posted by hosts in period 2 and the effort exerted in period 1. If hosts do not exert effort, guests realize that they have positive cost $c = k > 0$ and they do not exert effort in period 2: $\bar{\mu}_2(h^0) = 0$. Conversely, if hosts exert effort, then guests can only partially guess their cost of effort and $\bar{\mu}_2(h^1) > \pi$. Accordingly, exerting effort in period 1 rises hosts' prices $p_2^{pool}(h)$ and the probability to have a transaction $\alpha(\theta_2^{pool}(h))$. Still, hosts with histories h^0 are matched with guests with positive probability if $b > 0$.

As shown in Equation 2.1, the ex-ante guests' utility of a match does not vary across hosts with different histories, or, using a directed search term, different submarkets.

This is one of the main characteristics of the directed search framework and it is key to allow for the presence of hosts active in the market with different reputation levels and prices.

Corollary 1 (Guests Directed Search in Period 2)**.** *Guests' expected utility for a match in period 2 is the same across hosts active in the market:*

$$\begin{aligned}(a\bar{\mu}_2(h^1) + b - p_2^{pool}(h^1))\frac{\alpha(\theta_2^{pool}(h^1))}{\theta_2^{pool}(h^1)} &= U_2\\ (a\bar{\mu}_2(h^\emptyset) + b - p_2^\emptyset)\frac{\alpha(\theta_2^\emptyset)}{\theta_2^\emptyset} &\leq U_2\\ (a\bar{\mu}_2(h^0) + b - p_2^{pool}(h^0))\frac{\alpha(\theta_2^{pool}(h^0))}{\theta_2^{pool}(h^0)} &\leq U_2.\end{aligned}$$

Hosts with history h^1 are always matched with guests in period 2 with positive probability. Yet, hosts with histories $h^\emptyset$ and h^0 may not be matched if the expected guests' gross utility from a match with these hosts is too low. If this is the case, the last two conditions in Corollary 1 do not bind.

At the beginning of period 2, hosts can enter the market paying entry costs f. Once they enter, they will charge $p_2^\emptyset$ according to Proposition 1. In particular, the following entry condition characterizes the expected profits of new entrants:

$$p_2^\emptyset \alpha(\theta_2^\emptyset) \leq f. \tag{2.2}$$

Condition 2.2 is binding if a positive measure of hosts enters in period 2.

2.4 Hosts' Decisions: Period 1

In period 1, hosts who draw a positive cost of effort $c = k > 0$ choose whether to exert effort or not. Their decision is reported in their history and it changes the expected profits in period 2 according to Proposition 1.

Proposition 2 (Effort Decision in Period 1)**.** *In any reputational equilibrium, hosts who are matched with a guest in period 1 always exert effort if they have zero cost of effort, $c = 0$: $e_1(0) = 1$. If their cost of effort is positive, $c = k > 0$, they exert effort with probability $\omega \in [0, 1]$. ω is unique and it depends on the values of a, b, π, the cost of effort k, and the discount factor β.*

The proof of Proposition 2 is in Appendix A. Directly from the effort choice by hosts with

$c = k > 0$, it is possible to derive the guest's beliefs about hosts' effort in period 2.

Corollary 2 (Guests Beliefs Updating)**.** *Guests' interim beliefs about hosts' expected effort in period 2 are formed applying Bayes formula when possible:*

$$\begin{aligned} \bar{\mu}_2(h^1) &= \frac{\pi}{\pi + (1-\pi)\omega} \\ \bar{\mu}_2(h^\emptyset) &= \pi \\ \bar{\mu}_2(h) &= 0, \ \forall h \neq h^1, h^\emptyset. \end{aligned}$$

Guests (do not) update interim beliefs observing the price posted in period 2 (in equilibrium and off-equilibrium). In particular, given a history h, $\mu_2(h, p_2)$ is equal to $\bar{\mu}_2(h)$.

In period 1, hosts have not yet drawn their cost of effort when they post prices. Accordingly, the optimal pricing in period 1 is established in a condition of symmetric information between hosts and guests. Thus, guests' beliefs about hosts' effort in period 1 are not affected by prices: $\mu_1(p_1^\emptyset) = \pi + (1-\pi)\omega$. The optimal pricing is uniquely derived as follows.

Proposition 3 (Pooling in Prices in Period 1)**.** *In equilibrium, given guests' expected utility for a match U_1, hosts post prices $p_1^\emptyset$ and guests direct their search so as to form tightness $\theta_1^\emptyset$:*

$$\begin{aligned} \alpha'(\theta_1^\emptyset) &= \frac{U_1}{a(\pi + (1-\pi)\omega) + b - k(1-\pi)\omega + \beta\Delta\Pi} \\ p_1^\emptyset &= a(\pi + (1-\pi)\omega) + b - \frac{\theta_1^\emptyset}{\alpha(\theta_1^\emptyset)} U_1, \end{aligned}$$

where $\Delta\Pi$ represents the hosts' value of a transaction in terms of reputation updating. It is defined as follows:

$$\Delta\Pi = \Pi_2(a\bar{\mu}_2(h^1) + b)(\pi + (1-\pi)\omega) + (1-\pi)(1-\omega)\Pi_2(a\bar{\mu}_2(h^0) + b) - \Pi_2(a\bar{\mu}_2(h^\emptyset) + b).$$

If $a(\pi + (1-\pi)\omega) + b - k(1-\pi)\omega + \beta\Delta\Pi < U_1$, then $\theta_1^\emptyset = 0$ and $p_1^\emptyset = 0$.

The proof of this proposition is in Appendix A. Proposition 3 establishes that the value of a transaction in period 1 is not only related to the guests' expected utility $(a(\pi + (1-\pi)\omega) + b)$ and the cost of effort $(k(1-\pi)\omega)$, but it embeds a reputational value for hosts. If hosts draw $c = 0$, then they benefit from having a transaction since they get, with zero cost, expected profits $\Pi_2(a\bar{\mu}_2(h^1) + b)$ in period 2 with $\Pi_2(a\bar{\mu}_2(h^1) + b) \geq \Pi_2(a\bar{\mu}_2(h^\emptyset) + b)$. Conversely, if hosts draw $c = k > 0$, then having a transaction is not necessarily beneficial in terms of reputation updating. In particular, if $\omega = 0$, hosts with $c = k > 0$ get $\Pi_2(a\bar{\mu}_2(h^0) + b) \leq \Pi_2(a\bar{\mu}_2(h^\emptyset) + b)$.

In contrast with the multiple active submarkets in period 2, hosts do not show different histories in period 1. Thus, all guests direct their search to the only active submarket with expected utility U_1.

Corollary 3 (Guests Directed Search in Period 1)**.** *Guests' expected utility for a match in period 1 is defined as follows:*

$$(a(\pi + (1-\pi)\omega) + b - p_1^{\emptyset})\frac{\alpha(\theta_1^{\emptyset})}{\theta_1^{\emptyset}} = U_1.$$

Finally, at the beginning of period 1, hosts may enter the market paying entry costs f. The left-hand side of following entry condition characterizes the expected profits of entrants in period 1:

$$\begin{aligned} &(p_1^{\emptyset} - k(1-\pi)\omega)\alpha(\theta_1^{\emptyset}) \\ &+(1-\pi)(1-\omega)\beta\alpha(\theta_1^{\emptyset})p_2^{pool}(h^0)\alpha(\theta_2^{pool}(h^0)) \\ &+(\pi+(1-\pi)\omega)\beta\alpha(\theta_1^{\emptyset})p_2^{pool}(h^1)\alpha(\theta_2^{pool}(h^1)) \\ &+\beta(1-\alpha(\theta_1^{\emptyset}))p_2^{\emptyset}\alpha(\theta_2^{\emptyset}) \leq f. \end{aligned} \tag{2.3}$$

Condition 2.3 is binding if a positive measure of hosts enters in period 1.

In the remainder of this Section, I provide a formal definition of reputational equilibria and I analyze their existence and uniqueness.

Definition 1 (Reputational Equilibrium)**.** *A Reputational Equilibrium is defined by the following elements for period 1 and period 2, respectively:*

- $n_1, p_1, \mu_1(p_1), U_1, e_1(c, p_1)$: *the number of hosts who enter the market, the pricing decision, the guest' beliefs about hosts' effort, the guests' expected utility for a match, and the effort decision by hosts with cost of effort $c = 0$ and $c = k > 0$ in period 1, respectively;*
- $n_2(h), g_2(h, p_2(c,h)), p_2(c,h), \bar{\mu}_2(h), \mu_2(h, p_2(c,h)), U_2, e_2(c, p_2(c,h))$: *the number of hosts with history h present in the market, the number of guests who direct the search to hosts with certain history and price, the hosts' pricing decision for each cost and history, the guests' interim and updated beliefs about hosts' effort, the guests' expected utility for a match in period 2, and the effort decision by hosts with cost of effort $c = 0$ and $c = k > 0$ in period 2, respectively.*

The following conditions are satisfied in equilibrium:

1. *The market tightness in period 1 is defined as $\theta_1^{\emptyset} = \frac{1}{n_1}$;*
2. *The measures of hosts with history h^1 and h^0 in period 2 depend on the measure of hosts who entered in period 1, the probability of hosts drawing $c = 0$, π, and the probability to exert effort*

by hosts with $c = k > 0$, ω:

$$n_2(h^1) = (\omega(1-\pi) + \pi)\alpha(\theta_1^{\emptyset})n_1$$
$$n_2(h^0) = (1-\omega)(1-\pi)\alpha(\theta_1^{\emptyset})n_1;$$

3. *Guests in period 2 are assigned to different sets of hosts characterized by the couple formed by history and price. Tightness levels are such that $\theta_2(h) = \frac{g_2(h, p_2^{pool}(h))}{n_2(h)}$ and:*

$$\sum_h g_2(h, p_2^{pool}(h)) = G;$$

4. *The free-entry conditions 2.2 and 2.3 do not allow positive profits for hosts who enter the market in both periods;*
5. *Hosts post prices according to Propositions 1 and 3;*
6. *Guests' beliefs about hosts' effort in period 1 are $\mu_1(p) = \pi + (1-\pi)\omega$; whereas guests' beliefs in period 2 are formed according to Corollary 2;*
7. *Guests' expected utility levels from a match are defined according to Corollaries 1 and 3;*
8. *Hosts exert effort depending on their cost of effort according to Proposition 2 and Lemma 1.*

After having defined the equilibrium, I proceed with the theorem regarding its existence and uniqueness.

Theorem 1 (Existence and Uniqueness with Entry)**.** *If the measure of guests active in period 2 is greater than a threshold value $\bar{G}$, then reputational equilibrium exists and it is unique. In this equilibrium, a positive mass of hosts enters in both periods.*

The proof of Theorem 1 is in Appendix A.

2.5 Testable Predictions

Here I propose the main prediction of the model that can be directly tested using data from Airbnb. It follows from the comparison of two reputational equilibria with different entry costs for hosts.

Proposition 4 (Entry Costs and Effort)**.** *Consider two reputational equilibria in which the entry costs are f and f' with $f' > f$, and the measure of guests present in the market in period 2 is big enough to allow hosts' entry in both periods for f and f'. Then, in the reputational equilibrium associated with f', the probability that hosts with cost of effort $c = k > 0$ exert effort in period 1 is weakly higher than in the reputational equilibrium associated with f.*

Here I provide a heuristic proof for the proposition above.[8] If entry costs increase, the number of hosts who enter the market in period 2 decreases. The market is now tighter for guests in period 2 and guests' expected utility U_2 decreases. Conversely, hosts' expected profits increase and they increase more for hosts with better reputation. This is obvious if $b = 0$ and guests never direct their search to hosts with history h^0 in period 2. In this case, independently of the entry costs, hosts' expected profits in period 2 are zero if h^0. Conversely, the profits increase if hosts have histories h^1 or $h^\emptyset$. Thus, in period 1, hosts who draw $c = k > 0$ have stronger incentives to exert effort since the benefits of exerting effort - having a better reputation in period 2 - are higher. Accordingly, since more hosts with $c = k$ exert effort in period 1, the beliefs to have $c = 0$ with history h^1 drop. This leads to a lower premium of having good reputation.

The heuristics of the proof relies on the positive relationship between the tightness of the market is period 2 and the incentives to exert effort in period 1. In line with this mechanism, the empirical results in Section 5 address the effect of a change in competition, due to a variation in entry costs, over the effort exerted by hosts on Airbnb.

The identification strategy described in Section 4 proposes an instrumental variable that follows the channel highlighted in the proof of Proposition 4. Hosts anticipate the movement in tightness due to an exogenous change in entry costs. Thus, comparing hosts located in different areas, hosts exert more effort where the number of competitors drops more significantly: in a less competitive framework, exerting effort leads to greater reputational benefits.

In Section 6, two additional predictions are tested. They directly follow from the same comparative statics exercise of Proposition 4 and they can be tested using the same variations in entry cost.

Corollary 4 (Entry Costs, Profits and the Value of Reputation). *Consider two reputational equilibria in which the entry costs are f and f' with $f' > f$, and the measure of guests present in the market in period 2 is big enough to allow hosts' entry in both periods for f and f'. Then, in the reputational equilibrium associated with f':*

1. *Hosts' profits in period 2 are higher relative to the reputational equilibrium associated with f;*
2. *The value of reputation in period 2, that is the premium for having history h^1 is lower relative to the reputational equilibrium associated with f.*

[8] The interested reader may find the complete proof in Appendix A.

3 Empirical Setting and Dataset

In this Section, I introduce the empirical part of my work. First, I present the Airbnb setting. Then, I describe the regulation for short-term rentals in the city of San Francisco and I focus on the settlement agreement signed in May 2017 by the San Francisco City Council and Airbnb. Finally, I describe the unique dataset used for my analysis: I provide descriptive statistics about the population of Airbnb listings before and after the agreement signed in May 2017.

3.1 Airbnb

Airbnb is one of the leading digital platforms in the hospitality industry. It operates in more than 60,000 cities and it offers its members the possibility to arrange and offer lodging and other tourism experiences. Airbnb receives a commission fee for every transaction and it does not own any real estate listed on the platform. I restrict my analysis to lodging services and I denote the Airbnb members who arrange and offer accommodations as guests and hosts, respectively. To be an Airbnb member, a digital registration procedure is required. Airbnb guests need to provide personal information such as the email, and a phone number. The procedure to become an Airbnb host is different. It requires hosts to provide additional information and take photos of the listing; choose the days when they are willing to host; and set prices.[9] Further requirements are necessary for hosts due to local laws and regulations.

After being registered, guests and hosts appear on the Airbnb platform with a personal webpage. Guests can search for hosts that match the location and the period of their stay. Furthermore, other advanced filters are available to restrict the guests' search, such as price range and listings' characteristics. Guests can select hosts and visit their webpages. Then, they can choose to book the listing. If hosts accept guests' requests, their listings are officially booked.

After the guest's stay, host and guest have 14 days to review each other. Guests feedback consists of four elements:

1. A written comment;
2. Private comments to the host;
3. A one-to-five star rating about the overall experience;
4. Six specific ratings regarding the following categories:
   - The accuracy of the listing description;

[9]For more information regarding the registration procedure for Airbnb hosts, see the official Airbnb guide to becoming a host at www.airbnb.com/b/hosting_checklist.

- The check-in process at the beginning of the stay;
- The cleanliness of the listing;
- The communicativeness of the host;
- The listing location;
- The "value-for-money" of the stay.

Similarly, a host can review guests answering whether or not she would recommend the guest; writing a comment; and rating the guest considering the communicativeness, the cleanliness and how well the guest respected the rules of the house. Not all these elements are published on the platform and, for what concerns the guest feedback, only written comments are directly published on hosts' webpages. Ratings are not displayed singularly with the comments: only the rounded average of the score and subscores are published on the listing and the host webpages. In the same way, only the comment written by the host is published on the guest webpage.

3.2 Institutional Setting

Airbnb and other online marketplaces have had a sizable impact on the hospitality industry and many city councils have tried to regulate the rentals on digital platforms. I restrict my analysis to the city of San Francisco and I report here a synthetic chronology of the regulations adopted by the San Francisco City Council starting from the San Francisco Short-Term Rentals Regulation enacted in February 2015.

3.2.1 The San Francisco Short-Term Rentals Regulation (February 2015)

With an ordinance signed in October 2014 and effective from February 2015, the San Francisco city council legalized short-term rentals in the city. Before this ordinance, San Francisco banned short rentals in residential buildings. Rentals are considered "short-term" if the properties are rented for less than 30 consecutive nights at a time. Short-term rentals constitute the great majority of transactions occurring on hospitality digital platforms. Still, listings present on Airbnb can be exempt from the registration requirements if they only accept guests for periods of 30 or more days; or in case they are professional structures such as hotels and B&B. The regulation is mainly composed of the following parts:[10]

[10] For a comprehensive analysis of all the regulation's requirements, see the Short-Term Residential Rental Starter Kit provided by the San Francisco Office of Short-Term Rental at https://businessportal.sfgov.org/start/starter-kits/short-term-rental, and the official text of the ordinance at https://sfgov.legistar.com/View.

- Only San Francisco permanent residents who own or rent single-family dwellings in the city are eligible to engage in short-term rentals. In particular, hosts must reside in their dwellings for at least 275 days per year;
- Resident tenants must notify their landlords before engaging in short-term rentals;[11]
- Only the primary residence can be used for short-term rentals;
- When a host is absent, the dwelling can be rented for a maximum of 90 days per year;
- Hosts must obtain a permit and register at the Office of Short-Term Rental. Every two years, they must pay a $250 fee. Moreover, hosts are required to obtain a city business license;
- The San Francisco hotel tax must be collected from renters and paid to the cityFor Airbnb hosts, the platform automatically collects and pays such a tax for its hosts;
- Hosts must be covered by an insurance with a coverage of at least $500,000. Airbnb provides hosts with 1 million in coverage. Compliance to city building code requirements is necessary.

This regulation introduces several limitations on who can offer lodging service on Airbnb. To be legally present on the platform, hosts have to face additional costs and respect extra requirements.

In the first years after the introduction of the regulation, the enforcement of part of the law had proven to be difficult. In particular, regulators could not enforce the rules regarding hosts residence since registration rates at the Office of Short-Term Rental were very low and digital platforms did not disclose to the authorities any personal information regarding their hosts. Because of the difficulties regarding the enforcement of the law, San Francisco city council enacted an additional ordinance in June 2016 that required digital platforms to list on their websites only legal listings with official registration. Airbnb filed a suit against the City Council and, after a U.S. judge rejected the suit and postponed the enforcement of the new rules, an agreement was found in May 2017.

3.2.2 The Settlement Agreement with Airbnb (May 2017)

The agreement clarifies the role of digital platforms in the hosts registration process for short-term rentals. It has been signed, together with the San Francisco City Council, by Airbnb and another hospitality platform, HomeAway. The main resolutions are the following:[12]

[11] If the contract between tenant and landlord prohibits subletting, the landlord may evict the tenant. Moreover, tenants cannot charge more rent than they are paying to the landlord and rent control laws must be respected.

[12] All quotes are from the official announcement of the San Francisco City Attorney, available at https://www.sfcityattorney.org.

- From September 2017, new hosts willing to arrange a short-term rental on Airbnb or HomeAway have to "input their city Office of Short-Term Rental registration number (or application pending status) to post a listing";

- From September 2017, a "pass-through registration" system is implemented by Airbnb and HomeAway for hosts who are already registered on the platforms to send applications directly to the Office of Short-Term Rental for consideration. If the platforms receive notice of an invalid registration, they will cancel future stays and deactivate the listings;

- From January 2018, all hosts present on Airbnb and HomeAway are required to be registered. If some listings are not registered at this date, the platform will cancel future stays and deactivate the listings until a registration number (or application pending status) is provided.

3.3 InsideAirbnb Dataset

The dataset for this study comes from information on InsideAirbnb, a website that tracks all the Airbnb listings present in specific locations over time.[13] In my analysis, the dataset is formed by forty-seven snapshots of all the Airbnb listings present in San Francisco at forty-seven different dates from May 2015 to July 2019. Data scraping is performed at the beginning of each month with some months missing in 2015 and some multiple snapshots per month at the beginning of 2018.[14] I combine all the snapshots to form an unbalanced panel dataset composed of 30,266 listings and 350,099 listing observations over time. In each snapshot, listings are observed if they appear on the Airbnb website at the snapshot date. Accordingly, for each Airbnb listing in the dataset, entry, exit, and inactivity periods are identified.[15] When a listing is observed, several listing characteristics are displayed. Some are time-invariant such as the listing's location (longitude, latitude and neighborhood), and dwelling's characteristics. Some others update at each snapshot such as the number of guests' reviews and average star ratings, the price charged for one night at the snapshot date, the number of nights in which the listing is available after the snapshot, whether or not the listing displays the Office of Short-Term Rental registration number and whether the registration is necessary for the listing.

[13]All data are publicly available on InsideAirbnb. InsideAirbnb is "an independent, non-commercial set of tools and data that allows you to explore how Airbnb is really being used in cities around the world" and all scraped data are available under a Creative Commons CC0 1.0 Universal (CC0 1.0) license.

[14]The list of all snapshots follows: May 2015, September 2015, November 2015, December 2015, February 2016, April 2016, May 2016, June 2016, July 2016, August 2016, September 2016, October 2016, November 2016, December 2016, January 2017, February 2017, March 2017, April 2017, May 2017, June 2017, July 2017, August 2017, September 2017, October 2017, November 2017 (two snapshots), December 2017 (two snapshots), January 2018 (two snapshots), February 2018, March 2018, April 2018, May 2018, July 2018, August 2018, September 2018, October 2018, December 2018, January 2019, February 2019, March 019, April 2019, May 2019, June 2019, July 2019.

[15]Airbnb hosts can decide to remove their listings from Airbnb for a period of time and then re-enter with the same listing profile.

Descriptive statistics are reported in Table 1. Panel A presents the characteristics of all listings observed in the panel data from May 2015 to July 2019. All the reported variables correspond to the last snapshot in which listings are observed. The average amount of time that listings are present on Airbnb is approximately one year. The total number of reviews has a skewed distribution with more than half of listings having less than 5 reviews before exiting the platform. There is high variability in the price per night and the number of nights in which the listing is available after the snapshot, implying that performances on Airbnb widely vary across listings. In contrast, the variation of the average rating is much lower. The percentages regarding the number of hosts engaging in short-term rentals and displaying a registration number confirm two elements highlighted in the previous Section regarding Airbnb in San Francisco. First, short-term rentals constitute the great majority of transactions occurring on Airbnb (more than 80 percent). Second, before the Settlement Agreement, the regulation imposed by the San Francisco City Council was largely ignored. Panel B shows listings information regarding the number of reviews written between two consecutive snapshots and the averages of the ratings associated to these reviews. All the variables are constructed starting from the original variables shown in Panel A. I call these variables, the number of reviews per snapshot and the average ratings per snapshot. The number of reviews per snapshot is derived from the difference between the total number of reviews displayed in a snapshot and in the next one ($n_{i,t+1} - n_{i,t}$). Similarly, the average ratings per snapshot are computed using the average rating and the total number of reviews. I denote with $n_{i,t}$ and $\bar{R}^k_{i,t}$ the total number of reviews displayed for listing i at snapshot t and the average rating displayed for listing i at snapshot t for the category k, respectively. Then, the average rating per snapshot, $\bar{r}^k_{i,t}$, for listing i at snapshot t and category k where $k \in \{$*overall, accuracy, check-in, cleanliness, communication, location, value*$\}$ can be computed as follows:[16]

$$\bar{r}^k_{i,t} = \frac{\bar{R}^k_{i,t+1} n_{i,t+1} - \bar{R}^k_{i,t} n_{i,t}}{n_{i,t+1} - n_{i,t}}.$$

The number of reviews per snapshot varies by listing and snapshot. The average number of review per snapshot equals 1.5 with standard deviation 2.8. Much more limited variations are present for the average ratings per snapshot. The averages are higher than 9 for all the ratings with standard deviations always lower than 1.2. The average rating regarding the overall experience is 93.9 with standard deviation 9.2, that corresponds to an average of almost 5 stars with an extremely limited variation.[17]

[16]Since $\bar{R}^k_{i,t}$ are rounded averages, the procedure is likely to be affected by measurement errors. In order to reduce these errors, I drop the observations corresponding with values of $\bar{r}^k_{i,t}$ lower than 0 or greater than 10. For each rating, these values account for less than 2 percent of the sample. Moreover, I drop observations about snapshots with a number of reviews per snapshot greater than 26. I treat these snapshots as outliers due to the scraping method. They account for 0.08 percent of the sample.

[17]On the Airbnb platform guests can choose in a range of stars between 1 and 5. Still, the scraped variable

Table 1: Summary Statistics

	Mean	SD	N	Min	Max
Panel A					
Days in Airbnb	382.5	417.9	30,266.0	0.0	1,526.0
Total number of reviews	21.6	49.0	30,266.0	0.0	724.0
Percent of the listing population:					
Less than 5 reviews	58%	-	30,266.0	-	-
Between 5 and 10 reviews	9%	-	30,266.0	-	-
Between 10 and 20 reviews	10%	-	30,266.0	-	-
Between 20 and 50 reviews	11%	-	30,266.0	-	-
Between 50 and 100 reviews	6%	-	30,266.0	-	-
More than 100 reviews	6%	-	30,266.0	-	-
Short-term rentals	81%	-	30,266.0	-	-
Registration displayed or not required	42%	-	30,266.0	-	-
Price per night	206.8	189.2	30,009.0	0.0	1,500.0
Availability next 30 days	8.9	11.0	30,266.0	0.0	30.0
Availability next 60 days	20.1	22.3	30,266.0	0.0	60.0
Availability next 90 days	34.7	33.9	30,266.0	0.0	90.0
Minimum nights required	8.3	13.2	30,266.0	1.0	100.0
Panel B					
Average rating: overall	93.9	9.2	20,987.0	20.0	100.00
Number of reviews per snapshot	1.5	2.8	24,834.0	0.0	26.0
Average rating per snapshot: overall	93.3	8.8	14,849.0	0.0	100.0
Average rating per snapshot: accuracy	9.5	0.9	14,840.0	0.0	10.0
Average rating per snapshot: check-in	9.7	0.8	14,829.0	0.0	10.0
Average rating per snapshot: cleanliness	9.3	1.1	14,844.0	0.0	10.0
Average rating per snapshot: communication	9.6	0.8	14,840.0	0.0	10.0
Average rating per snapshot: location	9.4	0.9	14,828.0	0.0	10.0
Average rating per snapshot: value	9.1	1.0	14,827.0	0.0	10.0

Note: Panel A refers to every single listing present in the panel data combining the snapshots from May 2015 to July 2019. All the statistics refer to the last snapshot in which the listing is observed. The variable "Days in Airbnb" is derived considering the difference between the last and the first snapshot in which the listing is observed. The "Percent of the listing population" groups listings by the number of reviews that are displayed in their last snapshot. The variable "Price per night" presents the nominal prices charged by guests measured in US dollars. I drop few outliers reporting prices higher than \$1500. They account for 0.65 percent of the sample. Panel B refers to the variables constructed from the original dataset about the number of reviews written between two consecutive snapshots and the averages of the ratings associated to these reviews. Missing data regarding the variables "Average rating" are due to the high presence of listings with no reviews.

regarding the average rating for the overall experience varies from 0 to 100. All other scraped ratings varies from 0 to 10.

3.4 The Settlement Agreement: Exit, Entry, and Hosts' Selection

The Short-Term Rental Regulation has been effective since February 2015. However, as highlighted in Section 3.2, the enforcement of listings' registration at San Francisco Office of Short-Term Rental has proven to be difficult. The Settlement Agreement, effective from September 2017, addressed the enforcement difficulties of registration. It implemented a resolution that forced every eligible Airbnb listings to be registered before January 2018.

Figure 1 reports the percentage of Airbnb listings offering short-term lodging that displayed a registration number at each snapshot. Before September 2017, less than 15 percent of listings displays a registration number. Conversely, at the beginning of 2018, when the Settlement Agreement has been completely implemented, the percentage of listings offering short-term lodging with registration numbers reaches 100 percent and it stays constant afterwards.

Figures 2 and A.1 capture the change in the total number of Airbnb listings in San Francisco at each snapshot. Figure 2 shows the evolution of the number of Airbnb listings offering short-term lodging: from September 2016 until September 2017 the number of short-term listings remains constant between 8,000 and 9,000 units; then, when the "pass-through registration" system started to be at place, the number of listings sharply drops to 4,000 units in February 2018, when all eligible Airbnb units have to be registered. The number of short-term listings stays constant for the next months when the implementation of Settlement Agreement has been completed. To visualize the drop in the number of listings for different areas of San Francisco after the Settlement Agreement, Figures 3 and 4 present two maps with the location of Airbnb listings offering short-term lodging in San Francisco in September 2017 and in January 2018.

The evolution of the number of Airbnb listings that do *not* offer short-term lodging (from now on, long-term) is displayed in Figure A.1. The number of long-term listings, which are exempt from the regulation, steadily grows during the months in which the "pass-through registration" system starts to be implemented. Then, in August 2018, the number jumps from less than 1,000 units to more than 2,500 units in August 2018 and it continues to grow with more listings entering the platform without offering short-term lodging at the end of 2018 and at the beginning of 2019. Accordingly, the Settlement Agreement determined a selection in the type of listings that continued to be present on the platform after the implementation of the registration requirements.

In Table 2, I present some summary statistics to characterize this selection process. Listings are divided into four groups: Group A contains all listings that exit the platform before September 2017, when the implementation of the Settlement Agreement had not started yet. Group B contains all listings that enter the platform after September 2017. Listings in Group C enter the platform before September 2017 and exit after January 2018, when the implementation of the Settlement Agreement was completed.

Finally, Group D contains all listings that enter the platform before September 2017 and exit between September 2017 and January 2018. Accordingly, only listings belonging to Group C are present on Airbnb before and after the Settlement Agreement.

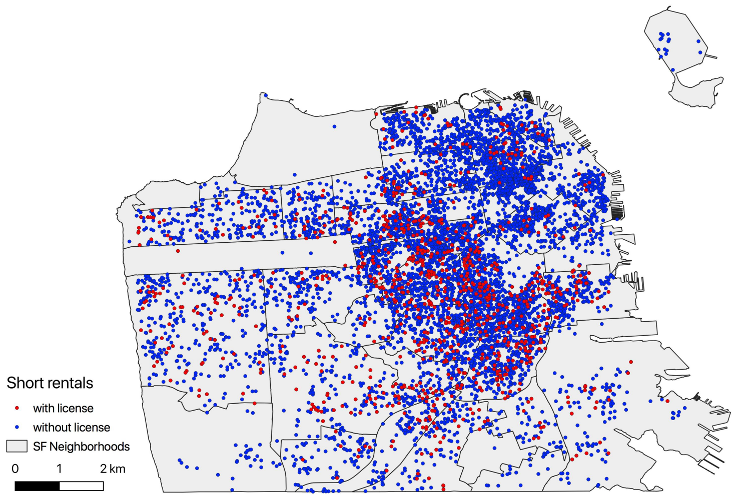
In Table 2, Panel A compares listings that are not affected by the Settlement Agreement (Group A) with listings that enter after September 2017 (Group B). This latter group of listings tends to engage in significantly longer rentals relative to Group A. In particular, hosts in Group B require guests to stay and rent their house for at least 15 consecutive nights, on average; whereas hosts in Group A require, on average, less than 4 consecutive nights. Accordingly, listings that enter after the Settlement Agreement are much less likely to engage in short-term rentals than those listings that are active before September 2017. The difference in the duration of the lodging services across groups may explain other differentials in terms of prices and the total number of reviews. The price per night charged by listings in Group A is significantly higher than the one charged by listings in Group B: shorter rentals tend to be more expensive. In addition, longer stays mechanically produce a lower stream of reviews over time. Moreover, listings in Group B tend to have significantly higher ratings than listings in Group A: this may be due to the different service duration, or to an improvement in the service quality provided by hosts.

A similar differential in the listing profiles is present in Panel B where listings that are present on Airbnb before and after the Settlement Agreement (Group C) are compared with those that enter before September 2017 and exit during the implementation of the new regulation (Group D). Survivors require guests to stay for more consecutive nights relative to listings in Group D and they charge lower prices. Still, they have a greater turnover since the number of reviews per snapshot is higher for Group C than Group D. Moreover, listings that stay after January 2018 have significantly higher ratings relative to those that exit before. In this sense, listings in Group C seem to be selected among those that present on Airbnb before the Settlement Agreement.

In Table A.1, I show statistics measured in September 2017 for listings in Groups C and D. In September 2017, listings in Group C have, on average, almost five times more reviews than in Group D, and enter the platform almost thirty days before. Listings in Group C charge significantly lower prices than listings in Group D and have higher ratings.

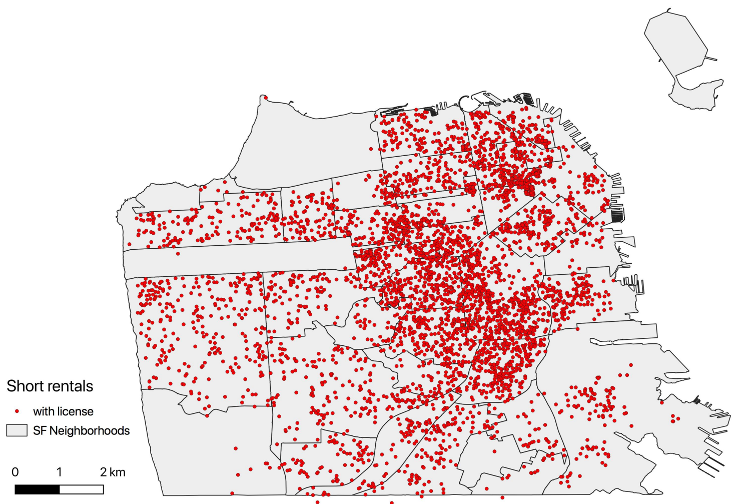
Accordingly, other than reducing the number of listings present on the platform, the regulation enforcement of the Settlement Agreement may have affected hosts' effort through the selection of hosts who stay after the enforcement of the registration. In order to tackle this issue, in Section 5, I restrict my analysis to those listings that enter before September 2017 and exit after January 2018, when the registration enforcement is completed (Group C). Still, external validity concerns may be at place restricting the sample on those listings that survive the registration enforcement. The estimated effects of competition over hosts' effort are based on a selected part of the population.

Figure 3: Location of Airbnb Short-Term Listings in San Francisco: September 2017



Note: The map shows the location of all Airbnb listings offering short-term lodging in San Francisco that were present for the snapshot associated with September 2017. Blue dots correspond to generic short-term Airbnb listings; whereas red dots correspond to listings that display a registration number.

Figure 4: Location of Airbnb Short-Term Listings in San Francisco: January 2018



Note: The map shows the location of all Airbnb listings offering short-term lodging in San Francisco that were present for the snapshot associated with January 2018. Red dots correspond to generic (registered) short-term Airbnb listings.

Table 2: Summary Statistics: the Settlement Agreement and Listings Selection

	Group A		Group B			
	Mean	SD	Mean	SD	Δ	$p-value$
Panel A						
Days in Airbnb	158.4	192.8	177.6	143.8	-19.2	0.0
Total number of reviews	11.0	24.8	7.5	16.7	3.5	0.0
Price per night	200.6	185.6	196.2	175.3	4.3	0.1
Availability next 30 days	11.1	11.5	8.9	10.9	2.2	0.0
Average rating per snapshot: overall	91.3	10.4	95.0	9.03	-3.6	0.0
Average rating per snapshot: accuracy	9.3	1.1	9.7	0.82	-0.3	0.0
Average rating per snapshot: check-in	9.5	0.9	9.8	0.78	-0.3	0.0
Average rating per snapshot: cleanliness	9.1	1.3	9.5	1.11	-0.4	0.0
Average rating per snapshot: communication	9.5	1.0	9.7	0.80	-0.2	0.0
Average rating per snapshot: location	9.3	1.1	9.6	0.8	-0.3	0.0
Average rating per snapshot: value-for-money	8.9	1.2	9.2	1.1	-0.3	0.0
Minimum nights required	3.8	7.2	19.0	17.8	-15.2	0.0
Short-term rentals	96%	-	45%	-	-0.5	-
Registration displayed or not required	34%	-	49%	-	-0.1	-
Number of listings	12,896	-	6,533	-	-	-
	Group C		**Group D**			
	Mean	SD	Mean	SD	Δ	$p-value$
Panel B						
Days in Airbnb	1,056.5	383.8	571.4	244.1	485.1	0.0
Total number of reviews	71.6	87.8	11.5	26.6	60.1	0.0
Price per night	207.5	177.0	244.0	230.7	-36.5	0.0
Availability next 30 days	7.02	9.7	4.9	9.7	2.1	0.0
Average rating per snapshot: overall	94.2	6.3	92.8	9.1	1.4	0.0
Average rating per snapshot: accuracy	9.6	0.7	9.5	1.0	0.2	0.0
Average rating per snapshot: check-in	9.8	0.5	9.7	0.8	0.1	0.0
Average rating per snapshot: cleanliness	9.5	0.8	9.2	1.2	0.3	0.0
Average rating per snapshot: communication	9.8	0.6	9.7	0.8	0.1	0.0
Average rating per snapshot: location	9.5	0.7	9.4	1.0	0.1	0.0
Average rating per snapshot: value-for-money	9.2	0.8	9.1	1.0	0.1	0.0
Minimum nights required	10.8	14.7	3.2	4.7	7.6	0.0
Short-term rentals	73%	-	99%	-	0.3	-
Registration displayed or not required	73%	-	5%		0.7	-
Number of listings	5,418	-	3,992	-	-	-

Note: The two panels of the table show and compare the profile of listings before and after the Settlement Agreement. All variables refer to the last snapshot in which the listing is observed apart from the variables "Average rating per snapshot". Listings are divided in four groups: Group A includes all listings that exit the platform before September 2017, when the implementation of the Settlement Agreement has not yet started. Group B includes all listings that enter the platform after September 2017. Group C includes all listings that enter the platform before September 2017 and exit after January 2018, when the implementation of the Settlement Agreement was completed. Group D includes all listings hat enter the platform before September 2017 and exit before January 2018. The last two columns provide the differences between the averages of relevant characteristics and the $p-value$ of the difference.

4 Identification Strategy

In this Section I discuss the identification strategy of the causal relationship between listing competition in Airbnb and the hosts' effort supporting the main prediction in Section 2. The model shows that variations in the entry costs change market tightness (the proportion of guests and hosts) affecting hosts' incentives to exert effort. In particular, if entry costs increase, hosts exert effort with higher probability to have better reputation in the future. The identification design closely follows the same channel: the variation in competition is due to the change in entry conditions established by the Settlement Agreement. The effect of this regulation on listing concentration varies across areas in San Francisco and the empirical strategy exploits these differences with a shock-based instrumental variable (IV) design.

The main estimating regression to capture the causal impact of competition on hosts' effort is the following:

$$\bar{r}_{i,t}^{effort} = \alpha_i + \rho_t + \beta \ln(L_{i,t}^j) + \varepsilon_{i,t}, \tag{4.1}$$

where α_i and ρ_t are the full set of dummy variables for each listing i and snapshot t. $\bar{r}_{i,t}^{effort}$ is a measure of hosts' effort. I use two rating categories as proxies for hosts' effort: check-in and communication. From now on, I denote the average rating per snapshot for listing i, snapshot t and category check-in and communication with $\bar{r}_{i,t}^{check}$ and $\bar{r}_{i,t}^{comm}$, respectively. I use $\bar{r}_{i,t}^{effort}$ to simultaneously refer to both average ratings. The focus on these two categories is justified by a principal component analysis performed on all the rating categories (average rating per snapshot). In Appendix C, Figure A.2 plots the loadings of all categories over the first two components. Check-in and communication are the most correlated ratings and their loadings separate from all others. In Section 6, I provide an estimation of the effort exerted by hosts using a control function approach to account for reviews' confounding factors related to guests' characteristics.

$L_{i,t}^j$ represents the degree of competition faced by listing i at snapshot t. It is defined as the sum of all listings offering short-term lodging at snapshot t within j kilometers of listing i. In my analysis I use three values for j: 0.5 kilometer, 1 kilometer, and 2 kilometers.[18] I use a logarithmic specification for the variable $L_{i,t}^j$ to capture non-linearities in the effect of competition.

With ordinary least squares (OLS), the correlation between $L_{i,t}^j$ and $\varepsilon_{i,t}$ produces inconsistent estimates of β. The main potential threat of endogeneity is with regard of the presence of omitted variables concerning the demand side. A high number of competitors is a signal of the attractiveness of the area and high demand. Thus, regressing $\bar{r}_{i,t}^{effort}$ over $\ln(L_{i,t}^j)$ may partially capture the impact of changes in demand over hosts' effort.

[18]These variables are created using information regarding latitude and longitude of each listing.

To tackle the endogeneity issues related to unobserved variations in the demand, I implement an IV strategy exploiting the Settlement Agreement between the San Francisco City Council and Airbnb. Accordingly, I restrict my analysis to listings offering short-term lodging that enter the platform before September 2017 and exit after January 2018 (Group C in Section 3.4).

In the same spirit as Dafny et al. (2012), Ashenfelter et al. (2015), and Chandra and Weinberg (2018), I propose a measure γ_i^j of the *predicted* change in the sum of listings within j kilometers of listing i due to the registration enforcement. The measure γ_i^j is the percentage of listings offering short-term lodging within j kilometers of listing i that display a registration number on their webpages few days before the Settlement Agreement became effective.[19] It is defined as follows:

$$\gamma_i^j = \frac{RL_{i,Sept2017}^j}{L_{i,Sept2017}^j},$$

where $RL_{i,Sept2017}^j$ and $L_{i,Sept2017}^j$ are the sum of listings offering short-term lodging with registration numbers and the total sum of listings offering short-term lodging, respectively, present at the beginning of September 2017 and within j kilometers of listing i. A value of γ_i^j close to 1 implies that the competition for listing i offering short-term lodging is not expected to change much since a high number of listings already displays a license. Conversely, low values of γ_i^j imply that the expected change in competition for listing i due to the Settlement Agreement is likely to be more relevant. Figure A.3 shows the distribution of γ_i^1 for the set of listings offering short-term lodging that enter the platform before September 2017 and exit after January 2018: more than 15 percent of Airbnb listings has no registered competitors within 1 kilometer in September 2017 ($\gamma_i^1 = 0$); whereas more than 60 percent of listings has at least 10 percent of registered competitors within 1 kilometer ($\gamma_i^1 > 0.1$).

The instrumental variable is formed by the product between γ_i^j and $post_{Nov2017}$: a dummy variable that takes value 1 for each snapshot after November 2017 and is zero otherwise.[20]

The power and the validity of this instrument depends on the strong correlation between γ_i^j and $L_{i,t}^j$ and on the assumption about the exclusion restriction. The "first stage" of the IV design documents a positive and significant relationship between the actual movement of the number of competitors for each listing and how the registration enforcement was expected to change the degree of competition.

[19] The snapshot in September 2017 was scrapped on September 2, 2017, whereas the new registration process started September 6, 2017. See http://www.sfexaminer.com/airbnb-launches-new-registration-system/.

[20] From Figure 2, November 2017 results to be the first snapshot with a significant drop in the number of listings offering short-term lodging in the platform.

The estimating equation of the "first stage" is the following:

$$\ln(L^j_{i,t}) = \alpha_i + \rho_t + \beta\gamma^j_i \times post_{Nov2017} + \varepsilon_{i,t} \tag{4.2}$$

where the endogenous variable $\ln(L^j_{i,t})$ is regressed over the expected change in competition due to the Settlement Agreement. Results with listings and snapshot fixed effects are in Table 3. The expected movement in the number of competitors, γ^j_i, is a good predictor for the actual change in competition occurring after November 2017: the higher is the value of γ^j_i , the lower is the expected negative effect of the Settlement Agreement over the hosts' population surrounding listing i. For each distance, all coefficients are positive and significant with a F-statistics much above the standard threshold to detect the presence of weak instruments.

To show further evidence of the predictive power of γ^j_i relative to the number of competitors for listing i over time, I illustrate the evolution of $L^j_{i,t}$ for different values of γ^j_i, and I integrate it with an event-study approach. I divide the population of Airbnb listings using the associated value of γ^1_i (the proportion of registered listings in September 2017 within 1 kilometer of listing i). In particular, I select those listings with a value of γ^1_i lower or equal than the 33th percentile of the distribution ($\gamma^1_i \leq 0.06$); and those with a value greater or equal than the 66th percentile of the distribution ($\gamma^1_i \geq 0.15$). Figure 5 shows the average values of $\ln(L^j_{i,t})$ over time for these two groups. The solid line depicts the evolution of the number of competitors for those listings that are predicted to be the most affected by the Settlement Agreement because of the low values of γ^1_i. Conversely, the dotted line shows the average value of $\ln(L^j_{i,t})$ for those listings that are predicted to be the least affected by the Settlement Agreement (high values of γ^1_i). From Figure 5 it is possible to observe that the drop of listing after the registration enforcement is much greater for the solid line relative to the dotted one, confirming the assumption that γ^1_i can predict the variation in the number of competitors surrounding each listing due to the Settlement Agreement.

I complement this analysis showing to which extent the dynamics in the number of competitors of each listing can be predicted by γ^j_i. I consider the following lead-lag model in which the degree of competition $L^j_{i,t}$ is regressed over the product between γ^j_i and a full set of dummy variables for each snapshot from September 2016 (one year before the registration enforcement started to be implemented) until January 2019 (one year after the end of the enforcement implementation):

$$\ln(L^j_{i,t}) = \alpha_i + \rho_t + \sum_{\tau=Sept2016}^{Jan2019} \beta_\tau\gamma^j_i \times 1(t=\tau) + \varepsilon_{i,t}. \tag{4.3}$$

I present the results of the OLS estimates of Equation 4.3 with an event study graph. I plot the estimated β_τ over the snapshot dates using the number of competitors within 1 kilometer in Figure 6. Before September 2017, the coefficients are close to zero and they do not exhibit a clear trend:

this evidence shows that the evolution of Airbnb listings before the Settlement Agreement is not correlated with γ_i^j. Conversely, the number of listings after September 2017 is positively correlated with γ_i^j: the number of listings offering short-term lodging sharply decreases after the Settlement Agreement (as it shown in Figure 2) and Airbnb listings are more likely to stay if the value of γ_i^j (the proportion of registered listings before the implementation of the Settlement Agreement) is higher.

With regard to the exclusion restriction, there is a list of arguments to support the assumption that the instrument ($\gamma_i^j \times post_{Nov2017}$) does not directly affect the dependent variable $\bar{r}_{i,t}^{effort}$, and it does only through its impact on the number of listings. There is no evidence that the San Francisco Short-Term Rental Regulation and the Settlement Agreement were motivated by policymakers' concerns over the quality of the services on hospitality platforms.[21]

To account for the selection of hosts who stay after the enforcement of the registration, I restrict my analysis to those listings that enter before September 2017 and exit after January 2018, when the registration enforcement is completed (Group C in Section 3.4). For this sample, the identification strategy excludes the presence of unobserved factors that affect hosts' effort and that are correlated with the predicted variations in the number of listings for different areas of San Francisco. Figures 5 and 6 present supportive evidence for this assumption: in Figure 5, the evolution of $\ln(L_{i,t}^j)$ for listings with different values of γ_i^j shows parallel trends before the Settlement Agreement. In line with this finding, Figure 6 shows that the estimated β_τ associated with the months before September 2017 are close to zero and no trend is detected. Accordingly, all evidence suggests that the instrumental variable is not correlated with unobservables affecting the evolution of the number of competitors.

In order to provide similar evidence regarding the correlation between the instrumental variable and unobservables affecting hosts' effort, I conduct the second event-study analysis.

I consider a lead-lag model in which the ratings regarding effort $\bar{r}_{i,t}^{effort}$ are regressed over the product between γ_i^j and a full set of dummy variables for each snapshot:

$$\bar{r}_{i,t}^{effort} = \alpha_i + \rho_t + \sum_{\tau=Sept2016}^{Jan2019} \beta_\tau \gamma_i^j \times 1(t=\tau) + \varepsilon_{i,t}. \tag{4.4}$$

As in the previous event-study, I present the results of the OLS estimates of Equation 4.4 with an event study graph. In Figures 7 and A.4, I plot the estimated β_τ over the snapshot dates considering the ratings regarding check-in and communication, respectively.

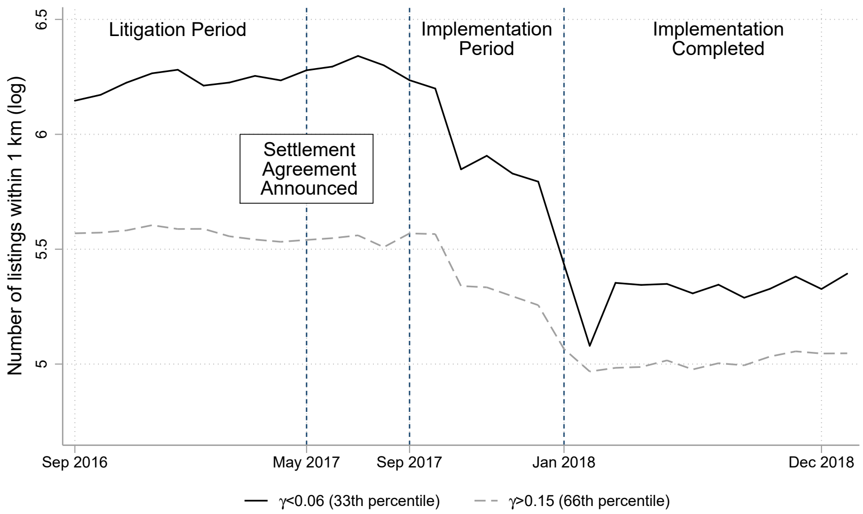
[21]The City Attorney, Dennis J. Herrera, never mentions the quality of the Airbnb service and the hosts' effort in his announcement of the Settlement Agreement, available at https://www.sfcityattorney.org.

Table 3: Impact of the Settlement Agreement on Competition (First Stage)

	$\ln(L_{i,t}^{0.5})$	$\ln(L_{i,t}^{1})$	$\ln(L_{i,t}^{2})$
$\gamma_i^{0.5} \times post_{Nov2017}$	1.364***		
	[0.0767]		
$\gamma_i^{1} \times post_{Nov2017}$		2.050***	
		[0.0771]	
$\gamma_i^{2} \times post_{Nov2017}$			2.477***
			[0.0694]
Constant	4.474***	5.741***	6.999***
	[0.00501]	[0.00358]	[0.00269]
Listing FE	✓	✓	✓
Snap FE	✓	✓	✓
Mean	4.226	5.488	6.737
F-test	648.6	1,342.6	2,932.8
R^2	0.103	0.112	0.132
N	57,274	57,274	57,274

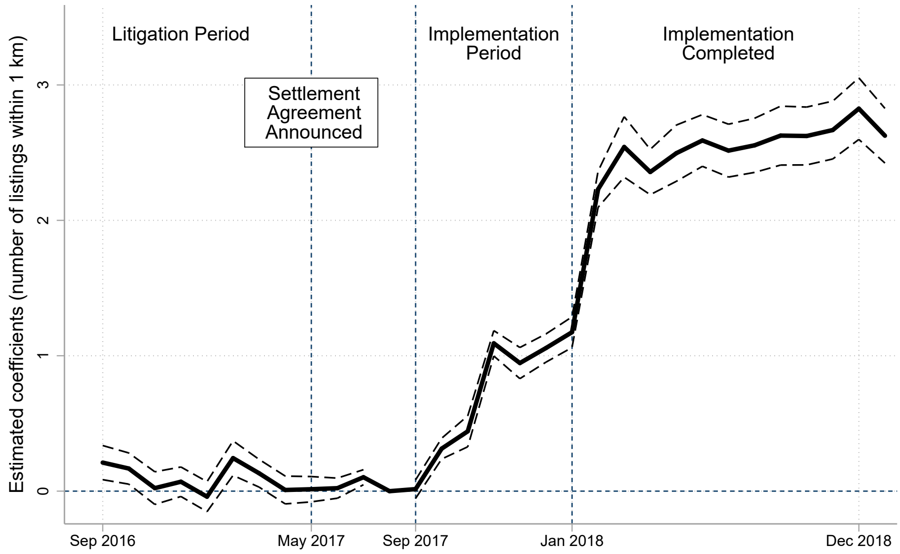
Note: Only listings offering short-term lodging that enter the platform before September 2017 and exit after January 2018 are considered. Standard errors clustered by listing are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure 5: Evolution of $\ln(L_{i,t}^{1})$ for Different Groups of Listings



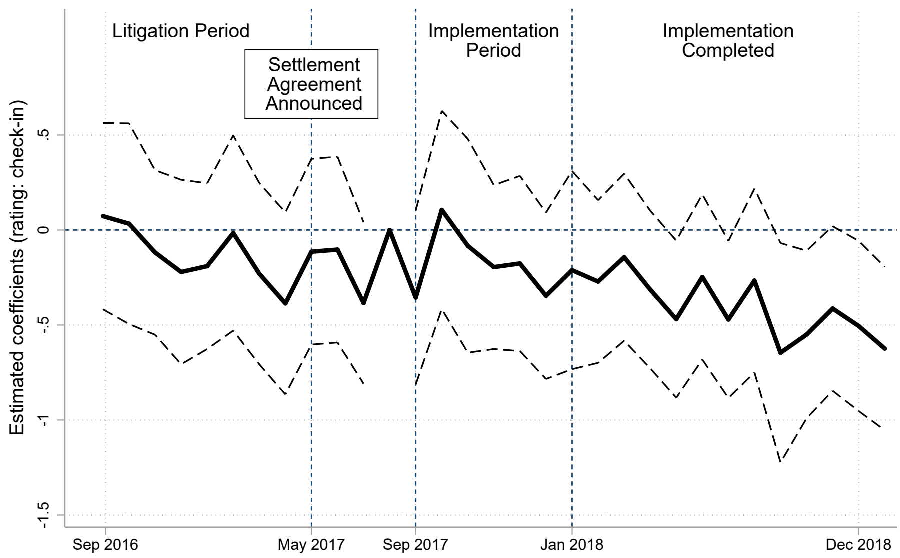
Note: The figure plots the average values of $\ln(L_{i,t}^{j})$ for two different groups of listings. The solid line represents the the average $\ln(L_{i,t}^{j})$ for those listings with $\gamma_i^1 \leq 0.06$; whereas the dotted line represents the the average $\ln(L_{i,t}^{j})$ for those listings with $\gamma_i^1 \geq 0.15$. Only listings offering short-term lodging that enter the platform before September 2017 and exit after January 2018 are considered. Listings represented by the solid line are predicted to be the most affected by the Settlement Agreement; whereas listings represented by the dotted line are predicted to be the least affected.

Figure 6: Estimated Coefficients from Equation 4.3



Note: In line with Equation 4.3, $\ln(L^1_{i,t})$ is regressed on listing and snapshot fixed effects, and on the products between γ^1_i and snapshot dummies. Standard errors are clustered by listing. The graph plots the estimated coefficients on these products. The value of the coefficient corresponding to August 2017 ($\hat{\beta}_{Aug2017}$) is normalized to zero. Only listings offering short-term lodging that enter the platform before September 2017 and exit after January 2018 are considered.

Figure 7: Estimated Coefficients from Equation 4.4: Ratings Regarding Check-in



Note: In line with Equation 4.4, $\bar{r}^{check-in}_{i,t}$ is regressed on listing and snapshot fixed effects, and on the products between γ^1_i and snapshot dummies. Standard errors are clustered by listing. The graph plots the estimated coefficients on these products. The value of the coefficient corresponding to August 2017 ($\hat{\beta}_{Aug2017}$) is normalized to zero. Only listings offering short-term lodging that enter the platform before September 2017 and exit after January 2018 are considered.

In both figures, the coefficients related to the months before September 2017 do not exhibit trends; whereas a negative trend is observable after the Settlement Agreement. These pieces of evidence confirm the exclusion restriction assumption: unobservables affecting ratings regarding hosts' effort do not correlate with the instrument variable before the registration enforcement. Furthermore, the negative trend after September 2017 supports the prediction of the model: when hosts face more competition (for higher values of γ_i^j), their incentives to exert effort decrease.

5 Main Results

I now present the main empirical results. To facilitate the comparison across different regressions, I always restrict my analysis to Airbnb listings that offer short-term lodging; enter the platform before September 2017; exit after January 2018; without missing data regarding the variables $\bar{r}_{i,t}^{check-in}$ and $\bar{r}_{i,t}^{comm}$. I consider data starting from September 2016 (one year before the registration enforcement started to be implemented) to January 2019 (one year after the end of the implementation). Throughout the next Sections, I allow the variance of residuals to differ across listings by clustering standard errors at listing level.[22]

I start by estimating the OLS panel regressions that relate hosts' effort to the degree of competition, as represented in Equation 4.1. Table 4 presents the results. For each rating, three regressions are performed: the independent variables vary depending on the distance used to delimit the competition faced by listings. The results suggest a not significant negative relationship between effort and competition measured as the sum of competitors within 0.5, 1 and 2 kilometers to each listing.

As described in Section 4, the OLS panel regressions are likely to be affected by the presence of omitted determinants of demand: the higher is the number of Airbnb hosts in a specific area, the greater is the area attractiveness for guests. Because of this, causality cannot be inferred from the OLS panel model. Accordingly, I take advantage of the variation in the degree of competition due to the Settlement Agreement to estimate the effect of competition over hosts' effort.

Table 3 shows the statistical and economic significance of γ_i^j in predicting the number of competitors faced by each listing after November 2017 (the "first stage"). In particular, when γ_i^j moves from 0 (no listing is registered before the Settlement) to 1 (all listing are registered), the number of competitors within 0.5 kilometers increases by more than 130 percent, those within 1

[22]The optimal clustering level should allow for correlations among competitor listings. Yet, the distance-based approach used to compute variables $L_{i,t}^j$ is unfeasible since listings have several competitors and should belong to different clusters simultaneously. It is tempting to consider geographically wider cluster levels, such as neighborhood. Still, only thirty-seven neighborhoods are present in San Francisco. Robust standard errors do not change coefficients' significance.

kilometer by more than 200 percent, and those within 2 kilometers by almost 250 percent. Before presenting the results regarding the IV estimates, I show the effect of the expected change in competition due to the regulation (the instrument) on the hosts' ratings about effort. The estimating equation presents the same functional form as Equation 4.2:

$$\bar{r}_{i,t}^{effort} = \alpha_i + \rho_t + \beta\gamma_i^j \times post_{Nov2017} + \varepsilon_{i,t} \tag{5.1}$$

This equation constitutes the "reduced form" of the IV estimates. Alternatively, it can be interpreted as a difference-in-difference design with a continuous control (the variable γ_i^j) that defines the extent to which the listing is affected by the regulation, i.e. the listing propensity to be treated by the shock.

Table 4: OLS Estimates of the Impact of Competition on Hosts' Effort

	$\bar{r}_{i,t}^{check-in}$			$\bar{r}_{i,t}^{comm}$		
	(1)	(2)	(3)	(4)	(5)	(6)
$\ln(L_{i,t}^{0.5})$	-0.0077			-0.0016		
	[0.012]			[0.011]		
$\ln(L_{i,t}^{1})$		-0.022			-0.012	
		[0.016]			[0.016]	
$\ln(L_{i,t}^{2})$			-0.036			-0.010
			[0.022]			[0.022]
Constant	9.876***	9.968***	10.09***	9.814***	9.880***	9.881***
	[0.0560]	[0.0947]	[0.155]	[0.0526]	[0.0945]	[0.157]
Listing FE	✓	✓	✓	✓	✓	✓
Snap FE	✓	✓	✓	✓	✓	✓
Mean	9.900	9.900	9.900	9.879	9.879	9.879
R^2	0.005	0.004	0.002	0.005	0.005	0.004
N	57,274	57,274	57,274	57,274	57,274	57,274

Note: Only listings offering short-term lodging that enter the platform before September 2017 and exit after January 2018 are considered. Standard errors clustered by listing are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Impact of the Settlement Agreement on Hosts' Effort (Reduced Form)

	$\bar{r}_{i,t}^{check-in}$			$\bar{r}_{i,t}^{comm}$		
	(1)	(2)	(3)	(4)	(5)	(6)
$\gamma_i^{0.5} \times post_{Nov2017}$	-0.155**			-0.153***		
	[0.0603]			[0.0564]		
$\gamma_i^{1} \times post_{Nov2017}$		-0.226***			-0.235***	
		[0.0846]			[0.0764]	
$\gamma_i^{2} \times post_{Nov2017}$			-0.244**			-0.224**
			[0.0978]			[0.0891]
Constant	9.841***	9.841***	9.841***	9.807***	9.807***	9.807***
	[0.0114]	[0.0114]	[0.0114]	[0.0113]	[0.0112]	[0.0112]
Listing FE	✓	✓	✓	✓	✓	✓
Snap FE	✓	✓	✓	✓	✓	✓
Mean	9.900	9.900	9.900	9.879	9.879	9.879
R^2	0.00250	0.00199	0.00219	0.00290	0.00225	0.00294
N	57,274	57,274	57,274	57,274	57,274	57,274

Note: Only listings offering short-term lodging that enter the platform before September 2017 and exit after January 2018 are considered. Standard errors clustered by listing are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: IV Estimates of the Impact of Competition on Hosts' Effort

	$\bar{r}_{i,t}^{check-in}$			$\bar{r}_{i,t}^{comm}$		
	(1)	(2)	(3)	(4)	(5)	(6)
$\ln(L_{i,t}^{0.5})$	-0.113**			-0.111***		
	[0.0447]			[0.0419]		
$\ln(L_{i,t}^{1})$		-0.110***			-0.115***	
		[0.0413]			[0.0375]	
$\ln(L_{i,t}^{2})$			-0.0986**			-0.0906**
			[0.0394]			[0.0360]
Constant	10.35***	10.47***	10.53***	10.31***	10.46***	10.44***
	[0.200]	[0.237]	[0.276]	[0.187]	[0.215]	[0.252]
Listing FE	✓	✓	✓	✓	✓	✓
Snap FE	✓	✓	✓	✓	✓	✓
Mean	9.900	9.900	9.900	9.879	9.879	9.879
R^2	0.00233	0.00111	0.000606	0.00232	0.000610	0.0000978
N	57,274	57,274	57,274	57,274	57,274	57,274

Note: Only listings offering short-term lodging that enter the platform before September 2017 and exit after January 2018 are considered. Standard errors clustered by listing are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 5 presents the results. For every specification and every rating, a negative and significant relationship between the instrument $\gamma_i^j \times post_{Nov2017}$ and hosts' effort is observed. Accordingly, lower values of γ_i^j, that predict a greater drop in the number of competitors for each listing, are associated with higher hosts' effort after November 2017. In this sense, a lower number of competitors is beneficial for hosts' ratings about effort. The parameters are also economically significant: when $\gamma_i^{0.5}$ changes from 0 to 1, $\bar{r}_{i,t}^{check-in}$ and $\bar{r}_{i,t}^{comm}$ decrease by almost 2 percent. Drops by more than 3 percent are associated with longer distances γ_i^1 and γ_i^2. It is important to note that the distributions of $\bar{r}_{i,t}^{effort}$ (presented in Table 2 for listings in Group C) are extremely concentrated and the magnitude of these changes roughly accounts for two third of standard deviation.

Finally, I turn to the IV estimates. $\ln(L_{i,t}^j)$ is the only endogenous variable in Equation 4.1, and only one instrumental variable is derived to predict the impact of the regulation, $\gamma_i^j \times post_{Nov2017}$. Then, the two-stage least squares parameters correspond to the ratio between the coefficients derived before for the "reduced form" and the "first stage" regressions (Equations 5.1 and 4.2, respectively).

The estimates are in Table 6. The results show a significant and negative effect of the number of competitors over hosts' effort in line with the parameters of the reduced form. The negative and significant impact of the IV is in contrast with the OLS estimates where the confounding factors due to demand side lead to inconclusive results. Moreover, the negative impact of the competition over hosts' effort is in line with the main prediction of the model (Proposition 2). In a less competitive setting, reputation concerns become more relevant and hosts exert effort with higher probability. In particular, a 1 percent decrease in the number of competitors leads to a increase of around 0.1 star for the ratings $\bar{r}_{i,t}^{check-in}$ and $\bar{r}_{i,t}^{comm}$. As commented before, the distributions of $\bar{r}_{i,t}^{effort}$ are very concentrated and a one-star change accounts for more than one standard deviation. Interestingly, the magnitude of the parameters monotonically decreases with the distance; the lowest parameters for both ratings are associated with a distance of 2 kilometers.

6 Extensions

In this Section, I present evidence supporting the other theoretical predictions proposed in Section 2. First, I show a negative causal relationship between the number of competitors and hosts' profits. To do so, I exploit the Settlement Agreement as exclusion restriction following the same empirical design explained in Section 4. Then, I analyze the monetary value of reputation and how it is affected by the change in competition due to the Settlement Agreement. My findings are in line with Elfenbein et al. (2015): they show that, in eBay, competition significantly increases the monetary value of reputation for eBay seller. Finally, I present evidence regarding the impact of the Settlement Agreement on Airbnb listings offering no short-term lodging services: a few months after the full implementation of the registration restriction, more than two thousand listings started to offer no short-term rentals. As a response to such variation, I observe that Airbnb hosts not offering short-term lodging decrease effort. This result confirms again the model's predictions: in Section 5, I study the relationship between effort and competition when the number of competitors decrease; whereas here I consider a positive shock in competition.

6.1 Competition and Profits

According to the model in Section 2, when the number of competitors decreases, hosts are more likely to exert effort and their expected profits to increase. Without information about the hosts' costs it is not possible to recover hosts' profits. Yet, I observe for each listing i and snapshot t the price charged per night, $p_{i,t}$, and the number of available nights to rent in the next 30 days, $available^{30}_{i,t}$. With these two variables, it is possible to compute a proxy regarding hosts' profits, $\pi^{30}_{i,t}$, as follows:[23]

$$\pi^{30}_{i,t} = p_{i,t}(30 - available^{30}_{i,t}).$$

I use the variable $\pi^{30}_{i,t}$ to study the relationship between profits and the number of competitors. The identification of the causal relationship follows the same strategy used for the ratings regarding hosts' effort. Accordingly, I conduct an event-study analysis to provide evidence about the correlation between the instrumental variable and unobservables affecting $\pi^{30}_{i,t}$. I consider a lead-lag model in which $\pi^{30}_{i,t}$ is regressed over the product between γ^j_i (the percentage of registered listings in September 2017 offering short-term lodging within j kilometers of listing i) and a full set of dummy variables for each snapshot:

$$\pi^{30}_{i,t} = \alpha_i + \rho_t + \sum_{\tau=Sept2016}^{Jan2019} \beta_\tau \gamma^j_i \times 1(t = \tau) + \varepsilon_{i,t}. \tag{6.1}$$

In Figure A.5 , I plot the estimated β_τ of Equation 6.1 over the snapshot dates. The coefficients related to the months before September 2017 do not exhibit trends (although the full set of dummy has not completely removed some seasonality effects). Conversely, the coefficients after January 2018 slightly decrease relative to the values before the Settlement Agreement. This is in line with the exclusion restriction assumption: unobservables affecting profits do not correlate with the instrument variable before the registration restriction's announcement. Furthermore, the negative trend after September 2017 shows that, hosts' revenues decrease when hosts face more competition,

[23] Measurement errors may affect this proxy variable. In particular, hosts may not be available to rent in some days for external reasons and not because the dwellings are already booked. Accordingly, the proxy may overestimate the total host's profits.

captured by higher values of γ_i^j. Similarly to the previous analysis regarding host's effort, I show the effect of the instrument $\gamma_i^j \times post_{Nov2017}$ on $\pi_{i,t}^{30}$, that is the "reduced form" of the IV estimates. The estimating equation is the following:

$$\pi_{i,t}^{30} = \alpha_i + \rho_t + \beta\gamma_i^j \times post_{Nov2017} + \varepsilon_{i,t}. \tag{6.2}$$

Table 7 presents the results. The coefficients regarding $\gamma_i^{0.5}$ and γ_i^1 show a negative and significant impact of the proportion of registered listings before the Settlement Agreement over hosts' revenues. Conversely, the coefficient of γ_i^2 is negative, but not significant. The effect is also economically relevant. When $\gamma_i^{0.5}$ varies from 0 to 1, hosts' monthly profits decrease by more than 400 US dollars after November 2017. Similarly, if γ_i^1 varies from 0 to 1, hosts' monthly profits decrease by more than 700 US dollars after November 2017. This findings support the model predictions. As before, the two-stage least squares coefficients are equal to the ratio between the parameters derived of the "reduced form" and the "first stage" regressions (Equations 6.2 and 4.2, respectively). Thus, the IV estimates provide the same result regarding the relationship between competition and profits. The estimates are in Table 8. The results show a significant and negative effect of the number of competitors within 0.5 and 1 kilometers over hosts' profits, whereas the the number of competitors within 2 have a negative, but not significant impact. The effect is also economically relevant: a 1 percent decrease in the number of competitors within 0.5 kilometer leads to a increase of more than 300 US dollars, that is almost 10%, in the monthly profits by hosts.

Table 7: Impact of the Settlement Agreement on Hosts' Profits (Reduced Form)

	$\pi_{i,t}^{30}$	$\pi_{i,t}^{30}$	$\pi_{i,t}^{30}$
$\gamma_i^{0.5} \times post_{Nov2017}$	-417.5*		
	[236.7]		
$\gamma_i^1 \times post_{Nov2017}$		-734.4**	
		[341.2]	
$\gamma_i^2 \times post_{Nov2017}$			-432.9
			[858.4]
Constant	4,140.2***	4,139.0***	4,138.7***
	[33.76]	[33.75]	[33.80]
Listing FE	✓	✓	✓
Snap FE	✓	✓	✓
Mean	3,804.6	3,804.6	3,804.6
R^2	0.0178	0.0181	0.0179
N	57,274	57,274	57,274

Note: Only listings offering short-term lodging that enter the platform before September 2017 and exit after January 2018 are considered. Standard errors clustered by listing are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 8: IV Estimates of the Impact of Competition on Hosts' Profits

	$\pi_{i,t}^{30}$	$\pi_{i,t}^{30}$	$\pi_{i,t}^{30}$
$\ln(L_{i,t}^{0.5})$	-306.4*		
	[173.5]		
$\ln(L_{i,t}^{1})$		-358.2**	
		[166.2]	
$\ln(L_{i,t}^{2})$			-174.8
			[346.6]
Constant	5,511.2***	6,195.8***	5,361.9**
	[776.3]	[951.6]	[2430.1]
Listing FE	✓	✓	✓
Snap FE	✓	✓	✓
Mean	3,804.6	3,804.6	3,804.6
R^2	0.00751	0.00617	0.0115
N	57,274	57,274	57,274

Note: Only listings offering short-term lodging that enter the platform before September 2017 and exit after January 2018 are considered. Standard errors clustered by listing are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

6.2 Competition and the Value of Reputation

The main prediction of the model regards the impact of competition on the reputational incentives to exert effort. In particular, I show that competition may erode the power of reputation to discipline users' behavior. In the previous Section, I report convincing evidence regarding Airbnb hosts' behavior in line with this prediction. However, using the same theoretical mechanism, it is possible to predict the effect on the economic value of having good reputation.

As it is pointed out in Section 2, the effect of having a history showing positive effort (good reputation) over the future expected profits depends on the proportion of hosts having such a history. If all hosts have good reputation, then histories lose their signaling power since guests cannot use them to update their beliefs about hosts' future effort decision. Conversely, when a smaller fraction of the hosts have a good reputation, then histories are signals of hosts' quality and they positively affect hosts' profits in the future. Accordingly, since the hosts' probability of exerting effort decreases with a higher degree of competition, it is possible to argue that the higher is the competition, the greater is the value of having good reputation.

This is in line with the findings by Elfenbein et al. (2015) observed using eBay data. They study the effect of quality certification (depending on users' ratings) on the probability to sell an item for eBay sellers. Their results show that the positive effect of certification is higher in more competitive settings and when certification is scarce.

Here I provide evidence supporting this prediction using Airbnb data. I use a hedonic regression approach in line with the literature that studies the value of reputation in online platforms (Cabral and Hortaçsu, 2010; Fan et al. (2016); Jolivet et al., 2016). Yet, I exploit the exogenous change in competition due to the Settlement Agreement to analyze how variations in the number of competitors affect the monetary value of reputation. To do this, I consider the following equation:

$$\pi_{i,t}^{30} = \alpha_i + \rho_t + \beta_1 \bar{R}_{i,t}^{effort} + \beta_2 \gamma_i^j \times post_{Nov2017} + \beta_3 \bar{R}_{i,t}^{effort} \times \gamma_i^j \times post_{Nov2017} + \varepsilon_{i,t}, \tag{6.3}$$

where $\pi_{i,t}$ is the proxy variable for hosts' profit defined above; and $\bar{R}_{i,t}^{effort}$ represents the average ratings regarding check-in and communication displayed for listing i at snapshot t. $\bar{R}_{i,t}^{effort}$ are the ratings displayed on the platform and, as they are averages, their variations over time are slower than (but in line with) the variations on the average ratings per snapshots $\bar{r}_{i,t}^{effort}$ that were analyzed in the previous Sections. Substituting $\bar{R}_{i,t}^{effort}$ with $\bar{r}_{i,t}^{effort}$ does not qualitatively alter the results of this Section.

The coefficient β_1 captures the relationship between the ratings displayed for listing i and its profits in the following thirty days. According to the mechanism presented above, more competitive settings should strengthen this relationship. To study this effect of competition, I multiply $\bar{R}_{i,t}^{effort}$ with the product $\gamma_i^j \times post_{Nov2017}$ that has a great predictive power over the changes in the number of competitors after the Settlement Agreement (see Section 4). In particular, higher values of γ_i^j predict a greater number of competitors staying in the market (since they have already complied with the regulation). In line with the model prediction, a greater amount of competitors should magnify the positive relationship between $\bar{R}_{i,t}^{effort}$ and $\pi_{i,t}$ resulting in a positive value for the coefficient β_3.

Results are in Table 9 for ratings regarding check-in and communication and for different distances. They support the prediction of a positive and significant impact of competition on the value of reputation ($\beta_3 > 0$). In particular, the effect of one-star increase in the average rating regarding check-in $\bar{R}_{i,t}^{check-in}$ over hosts' profit increases by more than $1,500$ US dollars if γ_i^j changes from 0 to 1. Similarly, the effect of one-star increase in the average rating regarding communication $\bar{R}_{i,t}^{comm}$ over hosts' profit increases by more than 600 US dollars if γ_i^j changes from 0 to 1.

6.3 Long-Term Listings and the Settlement Agreement

The Settlement Agreement has a profound impact on the enforcement of the Short-Term Rentals Regulation and, as a direct result, on the number of Airbnb listings offering short-term lodging active in San Francisco. However, it also has an indirect effect on the number of listings that do *not* rent short-term, but are present on the platform.

Table 9: Impact of the Settlement Agreement on the Value of Reputation

	$\pi_{i,t}^{30}$			$\pi_{i,t}^{30}$		
	(1)	(2)	(3)	(4)	(5)	(6)
$\gamma_i^{0.5} \times post_{Nov2017}$	-14,144.4***			-6,546.3*		
	[4,817.5]			[3,642.8]		
$\gamma_i^{1} \times post_{Nov2017}$		-18,496.2***			-7,702.1*	
		[5,474.9]			[4,076.2]	
$\gamma_i^{2} \times post_{Nov2017}$			-18,056.5***			-7,806.3*
			[5,540.3]			[4,322.5]
$\bar{R}_{i,t}^{check-in} \times \gamma_i^{0.5} \times post_{Nov2017}$	1,376.7***					
	[486.1]					
$\bar{R}_{i,t}^{check-in} \times \gamma_i^{1} \times post_{Nov2017}$		1,779.9***				
		[545.1]				
$\bar{R}_{i,t}^{check-in} \times \gamma_i^{2} \times post_{Nov2017}$			1,765.4***			
			[592.2]			
$\bar{R}_{i,t}^{comm} \times \gamma_i^{0.5} \times post_{Nov2017}$				615.9*		
				[370.3]		
$\bar{R}_{i,t}^{comm} \times \gamma_i^{1} \times post_{Nov2017}$					700.1*	
					[404.0]	
$\bar{R}_{i,t}^{comm} \times \gamma_i^{2} \times post_{Nov2017}$						741.1
						[483.0]
Constant	4,128.0***	4,259.9***	4,262.8***	3093.3***	3137.7***	3139.5***
	[579.6]	[576.5]	[578.9]	[665.0]	[670.3]	[670.9]
Listing FE	✓	✓	✓	✓	✓	✓
Snap FE	✓	✓	✓	✓	✓	✓
Mean	3,804.6	3,804.6	3,804.6	3,804.6	3,804.6	3,804.6
R^2	0.0181	0.0184	0.0183	0.0190	0.0193	0.0192
N	57,274	57,274	57,274	57,274	57,274	57,274

Note: Only listings offering short-term lodging that enter the platform before September 2017 and exit after January 2018 are considered. Standard errors clustered by listing are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

As shown in Figure A.1, the number of Airbnb listings that do not rent short-term (long-term) steadily grows from September 2017 to January 2018 and, after few months from the full implementation, the number jumps with an increase of more than two thousand units in August 2018.

The following regression provides convincing evidence that the massive entry in the market of long-term rentals is mainly due to hosts that could not comply with the Short-Term Rentals Regulation and exited the platform few months before. Entrants have new identification numbers and it is not possible to directly claim that they have already entered (and exited) the platform in previous periods. Yet, γ_i^j, the proportion of registered short-term listings in September 2017, has a significant predictive power over the variations of Airbnb long-term listings after the Settlement Agreement.

To observe this, I can repeat the "first stage" regression of the IV design with the following equation:

$$\ln(LL_{i,t}^j) = \alpha_i + \rho_t + \beta\gamma_i^j \times post_{Aug2018} + \varepsilon_{i,t} \tag{6.4}$$

where the endogenous variable $\ln(LL_{i,t}^j)$ represents the logarithm of the sum of all listings offering long-term lodging at snapshot t within j kilometers of listing i.

The dummy variable $post_{Aug2018}$ takes value 1 for each snapshot after August 2018 and it is zero otherwise.[24] I restrict my analysis on listings offering long-term lodging that enter the platform before September 2017 and exit after January 2018. This is the same period used for the previous analysis regarding short-term rentals. This restriction is necessary to remove the potential effect of selection on the estimates.

Results with listings and snapshot fixed effects are in Table 10 and they confirm that the expected movement in the number of competitors, γ_i^j , is a good predictor for the change in competition occurring after August 2018. The lower is the value of γ_i^j, the higher is the amount of listings that are likely to exit the market and, potentially, enter again offering long-term lodging. Thus, higher values of γ_i^j predict a negative effect of the Settlement Agreement over the hosts' population (renting long-term) surrounding listing i after August 2018. For each distance, all coefficients are negative and significant. As before, the F-statistics is above the standard threshold to detect weak instruments. Accordingly, Airbnb listings that rent long-term receive a positive shock in the number of competitors due to the Settlement Agreement; and this shock can be predicted using the same proportion γ_i^j used in the previous analysis. In addition, the exogeneity of γ_i^j seems reasonable since the registration restriction did not directly affect long-term Airbnb listings.

[24] From Figure A.1, August 2018 results to be the first snapshot with a significant jump in the number of listings offering long-term lodging in the platform.

Table 10: Impact of the Settlement Agreement on Competition (First Stage)

	$\ln(LL^{0.5}_{i,t})$	$\ln(LL^{1}_{i,t})$	$\ln(LL^{2}_{i,t})$
$\gamma_i^{0.5} \times post_{Aug2018}$	-2.351***		
	[0.409]		
$\gamma_i^{1} \times post_{Aug2018}$		-3.147***	
		[0.375]	
$\gamma_i^{2} \times post_{Aug2018}$			-3.067***
			[0.228]
Constant	1.685***	2.898***	4.146***
	[0.113]	[0.0727]	[0.0519]
Listing FE	✓	✓	✓
Snap FE	✓	✓	✓
Mean	2.633	3.812	5.041
F-test	84.49	226.5	800.0
R^2	0.515	0.534	0.579
N	3,326	3,326	3,326

Note: Only listings offering long-term lodging that enter the platform before September 2017 and exit after January 2018 are considered. Standard errors clustered by listing are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 11: IV Estimates of the Impact of Competition on Hosts' Effort

	$\bar{r}^{check-in}_{i,t}$			$\bar{r}^{comm}_{i,t}$		
	(1)	(2)	(3)	(4)	(5)	(6)
$\ln(LL^{0.5}_{i,t})$	-0.154			-0.259		
	[0.0988]			[0.160]		
$\ln(LL^{1}_{i,t})$		-0.185**			-0.193	
		[0.0940]			[0.131]	
$\ln(LL^{2}_{i,t})$			-0.239**			-0.270*
			[0.111]			[0.157]
Constant	10.20***	10.47***	10.93***	10.39***	10.50***	11.06***
	[0.204]	[0.304]	[0.485]	[0.338]	[0.449]	[0.712]
Listing FE	✓	✓	✓	✓	✓	✓
Snap FE	✓	✓	✓	✓	✓	✓
Mean	9.818	9.818	9.818	9.826	9.826	9.826
R^2	0.00260	0.000138	0.00118	0.00145	0.00000842	0.000829
N	3,326	3,326	3,326	3,326	3,326	3,326

Note: Only listings offering long-term lodging that enter the platform before September 2017 and exit after January 2018 are considered. Standard errors clustered by listing are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Therefore, it is possible to use this opposite shock in the number of competitors to test again the main prediction of the model. In this case, the increase in the number of competitors should have a negative impact on the hosts' incentives to exert effort.

Table 11 shows the results of the IV estimates regarding the impact of $\ln(LL^j_{i,t})$ on $\bar{r}^{effort}_{i,t}$ for listings offering no short-term lodging. In line with the previous findings, the results confirm a negative effect of the number of competitors over hosts' effort. I restrict my analysis over listings that do not rent short-term and that were already present on the platform before the Settlement Agreement. This restriction is necessary to avoid selection effects that may be at work after the regulation enforcement. Still, it also bring the focus on a small sample of listings reducing the significance of the parameters. All coefficients are negative and their magnitude is similar to one presented in Table 6 relative to short-term listings. Yet, only three out of six coefficients are statistically significant. In particular, the coefficients regarding the $\ln(LL^2_{i,t})$ are the most significant and negative.

7 Robustness Checks

Here I examine the robustness of the IV estimates presented in Section 5. I provide an estimation of hosts' effort and I show the negative impact of the competition on the estimated hosts' effort.

7.1 Effort Estimation

Submitting reviews, guests answer several questions about their stay. Many dimensions of the lodging service are part of the guests' feedback, and not all regards the effort exerted by hosts during the stay. In Section 4, two rating categories are used as proxies for hosts' effort: check-in and communication. Still, although guests' feedback may be informative about hosts' effort, reviews are also affected by other factors related to guests' characteristics. To account for such confounding factors, I provide here a hosts' effort estimation using a control function approach. I denote with $\bar{r}^{location}_{i,t}$ the average rating per snapshot for listing i, snapshot t and the category location. Taking advantage of the fact that location should not depend on hosts' effort, in contrast with check-in and communication, I propose the following functional forms:

$$\bar{r}^{location}_{i,t} = \theta_i + guest^{location}_{i,t} \tag{7.1}$$

$$\bar{r}^{effort}_{i,t} = e_{i,t} + guest^{effort}_{i,t}, \tag{7.2}$$

where θ_i is the fixed quality of listing i; $e_{i,t}$ is the effort exerted by the host of listing i at snapshot t; $guest^{location}_{i,t}$ and $guest^{effort}_{i,t}$ account for the guests' specific characteristics about the location and effort such as attitude, tastes or generosity. The control function approach relies on the following equation:

$$guest_{i,t}^{effort} = \alpha + \beta guest_{i,t}^{location} + \epsilon_{i,t}, \tag{7.3}$$

with $E(guest_{i,t}^{location}\epsilon_{i,t}) = 0$ and $\beta \neq 0$. Equation 7.3 assumes a common linear relationship between guests' characteristics for all ratings in the dataset. It allows guests to have different values of $guest_{i,t}^{location}$ and $guest_{i,t}^{effort}$, but a common linear relationship is always present for every guest up to the orthogonal error $\epsilon_{i,t}$.[25] Plugging Equation 7.3 into the previous system of equations, I derive the following fixed effect panel regression:

$$\begin{aligned}
\bar{r}_{i,t}^{effort} &= e_{i,t} + guest_{i,t}^{effort} \\
&= e_{i,t} + \alpha + \beta guest_{i,t}^{location} + \epsilon_{i,t} \\
&= e_{i,t} + \alpha + \beta(\bar{r}_{i,t}^{location} - \theta_i) + \epsilon_{i,t} \\
\bar{r}_{i,t}^{effort} &= \alpha - \beta\theta_i + \beta\bar{r}_{i,t}^{location} + e_{i,t} + \epsilon_{i,t}.
\end{aligned} \tag{7.4}$$

In Equation 7.4, $\bar{r}_{i,t}^{effort}$ is regressed on $\bar{r}_{i,t}^{location}$ with a constant and a listing fixed effect accounting for $\alpha - \beta\theta_i$. Accordingly, the host effort $e_{i,t}$ can be estimated from the residuals of fixed effect panel regression with noise $\epsilon_{i,t}$. To have a consistent estimate of β (and unbiased measures of effort), the following orthogonality conditions need to hold:

$$E[\bar{r}_{i,t}^{location}\epsilon_{i,t}|\theta_i] = 0 \tag{OC_1}$$

$$E[\bar{r}_{i,t}^{location}e_{i,t}|\theta_i] = 0. \tag{OC_2}$$

Condition OC_1 directly follows from the assumption 7.3 and the orthogonality of the error $\epsilon_{i,t}$ with $guest_{i,t}^{location}$. Differently, condition OC_2 imposes hosts' effort to not be correlated with deviations of $\bar{r}_{i,t}^{location}$ from the fixed quality θ_i.

I provide empirical evidence supporting condition OC_2 studying the relationship between the effort measures $e_{i,t}^{check}$, $e_{i,t}^{comm}$, the location rating $\bar{r}_{i,t}^{location}$ and a different proxy for hosts' effort present in the dataset: hosts' response rate. This variable represents the percentage of new inquiries or lodging requests to which the host responded within 24 hours in the past 30 days before each snapshot.[26] In case of hosts with multiple listings, the variable does not adjust and it considers all new inquires received by a host. To account for this, I restrict the analysis to single listings, i.e.

[25]The common relationship can be relaxed allowing the parameter β to change over time-invariant group of listings. Still, all results presented in this Section do not qualitatively change when I allow for different values of β with a random coefficient approach (see Appendix E.1).

[26]For more information regarding how the response time is computed, see the official Airbnb webpage at www.airbnb.com/help/article/430/what-is-response-rate-and-how-is-it-calculated.

listings whose hosts do not manage multiple properties on Airbnb.[27] I regress the variable hosts' response rate over the effort measures $e_{i,t}^{check}$, $e_{i,t}^{comm}$, and the location rating $\bar{r}_{i,t}^{location}$ controlling for listing fixed effects (see Table 12).[28] The results support condition OC_2: hosts' response rate is not significantly correlated with deviations of $\bar{r}_{i,t}^{location}$, whereas it is positively and significantly correlated with the effort dimensions.

In the remaining part of the Section, I study hosts' effort showing the previous results using the effort measures $e_{i,t}^{check}$, $e_{i,t}^{comm}$ estimated as residuals of the regression in Equation 7.4. The identification strategy presented in Section 4 can be replicated using $e_{i,t}^{check}$, $e_{i,t}^{comm}$ as proxies for hosts' effort.[29] Table A.2 presents the OLS panel regressions of hosts' effort and the number of competitors as shown in Equation 4.1. These regressions show not significant results, similar to the case of ratings $\bar{r}_{i,t}^{effort}$ (Table 4). Demand-driven confounding factors are likely at place and endogeneity issues affect the regressions' coefficients. Thus, I estimate the effect of competition over hosts' effort considering variations in the number of competitors due to the Settlement Agreement.

First, I present results about the "reduced form" of the IV estimates considering the functional form of Equation 5.1. Table 13 shows the results.

The negative relationship between the instrument $\gamma_i^j \times post_{Nov2017}$ and hosts' effort holds, and it is statistically significant for both measures $e_{i,t}^{check}$ and $e_{i,t}^{comm}$. The economic significance of the relationship is also present. When $\gamma_i^{0.5}$ changes from 0 to 1, $e_{i,t}^{check}$ and $e_{i,t}^{comm}$ decrease by almost 0.2 units; and longer distances, γ_i^1 and γ_i^2, are associated with drops greater than 0.2 units. It is important to recall that, because of its nature of residuals, the measure has a zero sample mean with standard deviation equal to 0.47. Thus, a change of 0.2 accounts for almost one half of standard deviation.

Similar results characterize the IV estimates, presented in Table 14. A negative relationship between the number of competitors and hosts' effort is present for $e_{i,t}^{check}$ and $e_{i,t}^{comm}$. Accordingly, the negative relationship between number of competitors and hosts' effort holds even after removing confounding factors due to guests' characteristics.

[27] At each snapshot I observe listing and host identification numbers. Single listings constitute the 48 percent of total amount of Airbnb listings in the dataset.

[28] The variable hosts' response rate takes values from 0 to 1. A higher percentage corresponds to a faster rate of host's replies.

[29] I use information regarding all listings in the dataset to compute the effort measures by Equation 7.4. To replicate the previous analysis I consider the same restrictions as in Section 5. The number of observations are slightly different relative to the previous analysis since I have to exclude listings with missing information about $\bar{r}_{i,t}^{location}$ to estimate the effort measures.

Table 12: Evidence Supporting Assumption OC_2: Response Rate, $\bar{r}_{i,t}^{location}$, $e_{i,t}$

	Response rate	Response rate	Response rate
$\bar{r}_{i,t}^{location} \times 100$	0.0661		
	[0.0499]		
$e_{i,t}^{comm} \times 100$		0.216**	
		[0.106]	
$e_{i,t}^{check} \times 100$			0.314***
			[0.112]
Constant	0.967***	0.974***	0.974***
	[0.00478]	[0.000000118]	[0.000000405]
Listing FE	✓	✓	✓
Mean	0.972	0.972	0.972
R-squared	.0000752	.0002651	.0001957
N	68,371	68,371	68,371

Note: Only listings offering short-term lodging that enter the platform before September 2017 and exit after January 2018 are considered. Standard errors clustered by listing are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 13: Impact of the Settlement Agreement on Hosts' Effort (Reduced Form)

	$e_{i,t}^{check-in}$			$e_{i,t}^{comm}$		
	(1)	(2)	(3)	(4)	(5)	(6)
$\gamma_i^{0.5} \times post_{Nov2017}$	-0.146**			-0.117**		
	[0.0610]			[0.0581]		
$\gamma_i^{1} \times post_{Nov2017}$		-0.212**			-0.204***	
		[0.0848]			[0.0766]	
$\gamma_i^{2} \times post_{Nov2017}$			-0.214**			-0.181**
			[0.0981]			[0.0899]
Constant	-0.0441***	-0.0440***	-0.0440***	-0.0563***	-0.0563***	-0.0562***
	[0.0117]	[0.0117]	[0.0117]	[0.0114]	[0.0114]	[0.0114]
	[0.0116]	[0.0116]	[0.0116]	[0.0114]	[0.0114]	[0.0114]
Listing FE	✓	✓	✓	✓	✓	✓
Snap FE	✓	✓	✓	✓	✓	✓
Mean	0.006	0.006	0.006	0.007	0.007	0.007
R^2	0.00210	0.00211	0.00212	0.00235	0.00237	0.00237
N	55,587	55,587	55,587	55,587	55,587	55,587

Note: Only listings offering short-term lodging that enter the platform before September 2017 and exit after January 2018 are considered. Standard errors clustered by listing are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 14: IV Estimates of the Impact of Competition on Hosts' Effort

	$e_{i,t}^{check-in}$			$e_{i,t}^{comm}$		
	(1)	(2)	(3)	(4)	(5)	(6)
$\ln(L_{i,t}^{0.5})$ ×10	-0.970**			-0.839**		
	[0.437]			[0.412]		
$\ln(L_{i,t}^{1})$ ×10		-0.949**			-0.873**	
		[0.405]			[0.365]	
$\ln(L_{i,t}^{2})$ ×10			-0.832**			-0.669*
			[0.386]			[0.351]
Constant	0.391**	0.502**	0.539**	0.319*	0.446**	0.412*
	[0.196]	[0.233]	[0.271]	[0.184]	[0.209]	[0.246]
Listing FE	✓	✓	✓	✓	✓	✓
Snap FE	✓	✓	✓	✓	✓	✓
Mean	0.006	0.006	0.006	0.007	0.007	0.007
R^2	0.000110	0.000149	0.000311	0.0000956	0.000138	0.000375
N	55,587	55,587	55,587	55,587	55,587	55,587

Note: Only listings offering short-term lodging that enter the platform before September 2017 and exit after January 2018 are considered. Standard errors clustered by listing are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

8 Conclusion

In this work, I provide theoretical and empirical evidence regarding the negative effect on the incentives to exert effort of the number of competitors using a model of reputation concerns. First, I develop a reputation model in a directed search framework where movements in entry costs impact the number of hosts in the market and their incentives to exert effort. Then, using a unique dataset of Airbnb, I identify the causal relationship between the number of competitors on the platform and hosts' effort. To do so, I consider a change in the regulation regarding the registration enforcement of Airbnb hosts in San Francisco in September 2017. I obtain a negative and significant effect regarding the extent of competition over hosts' effort. All empirical results are in line with the main predictions of the model.

The main limitation of my work regards the structure of the dataset and the available pieces of information concerning transactions and effort. All the proxies that I use to estimate hosts' effort are extracted from the Airbnb feedback system. In this sense, my analysis considers only the hosts' effort exerted in reviewed transactions.

From a policy perspective, the results of my work suggest that limiting the number of competitors in a platform increases the profits of those agents who remain and it may be beneficial in terms of services' quality. In addition to hosts' (positive) selection, hosts have stronger incentives to exert effort and provide good quality services.

Accordingly, rental restrictions, such as the San Francisco Short-Term Rentals Regulation, favor local hosts complying with the regulatory terms without undermining hosts' quality provision. Moreover, this work sheds light on a trade-off between quantity and quality of transactions in the context of platform design. Several platforms (Airbnb included) charge a percent fee on the total price of each transaction between agents. Therefore they have incentives to lower entry costs, attract more users and foster more exchanges. Still, my work shows that an increase in entry costs leads hosts to charge higher prices and exert more effort. Transactions' quality increases as well as platform's profit per transaction. Thus, the total effect of an increase in entry costs on platforms' profit is ambiguous. In line with these policy implications, further research is necessary to investigate the optimal entry fee for the efficiency of the market and for platforms.

References

Aghion, P., Bloom, N., Blundell, R., Griffith, R., and Howitt, P. (2005). Competition and innovation: An inverted-u relationship. *The Quarterly Journal of Economics*, 120(2):701–728.

Aghion, P., Harris, C., Howitt, P., and Vickers, J. (2001). Competition, imitation and growth. *Review of Economic Studies*, 68(3):467–492.

Ashenfelter, O. C., Hosken, D. S., and Weinberg, M. C. (2015). Efficiencies brewed: Pricing and consolidation in the US beer industry. *RAND Journal of Economics*, 46(2):328–361.

Bar-Isaac, H. (2005). Imperfect competition and reputational commitment. *Economics Letters*, 89(2):167–173.

Bar-Isaac, H. and Tadelis, S. (2008). Seller Reputation. *Foundations and Trends® in Microeconomics*, 4(4):273–351.

Burdett, K., Coles, M., Kiyotaki, N., and Wright, R. (1995). Buyers and sellers: should I stay or should I go? *American Economic Review*, 85(2):281–286.

Burdett, K., Shi, S., and Wright, R. (2001). Pricing and Matching with Frictions. *Journal of Political Economy*, 109(5):1060–1085.

Cabral, L. (2012). Reputation on the internet. *The Oxford handbook of the digital economy*, pages 343–354.

Cabral, L. and Hortaçsu, A. (2010). The dynamics of seller reputation: Evidence from ebay. *Journal of Industrial Economics*, 58(1):54–78.

Chandra, A. and Weinberg, M. (2018). How does advertising depend on competition? evidence from us brewing. *Management Science*, 64(11):5132–5148.

Cho, I.-K. and Kreps, D. M. (1987). Signaling Games and Stable Equilibria. *The Quarterly Journal of Economics*, 102(2):179–221.

Dafny, L., Duggan, M., and Ramanarayanan, S. (2012). Paying a premium on your premium? consolidation in the us health insurance industry. *American Economic Review*, 102(2):1161–85.

Delacroix, A. and Shi, S. (2013). Pricing and signaling with frictions. *Journal of Economic Theory*, 148(4):1301–1332.

Elfenbein, D. W., Fisman, R., and McManus, B. (2015). Market structure, reputation, and the value of quality certification. *American Economic Journal: Microeconomics*, 7(4):83–108.

Fan, Y., Ju, J., and Xiao, M. (2016). Reputation premium and reputation management: Evidence from the largest e-commerce platform in China. *International Journal of Industrial Organization*, 46:63–76.

Farronato, C. and Fradkin, A. (2018). The welfare effects of peer entry in the accommodation market: The case of airbnb. Technical report, National Bureau of Economic Research.

Fradkin, A. (2015). Search frictions and the design of online marketplaces. *Work. Pap., Mass. Inst. Technol.*

Fradkin, A. (2017). Search, matching, and the role of digital marketplace design in enabling trade: Evidence from airbnb. *Matching, and the Role of Digital Marketplace Design in Enabling Trade: Evidence from Airbnb (March 21, 2017).*

Horton, J. J. (2019). Buyer uncertainty about seller capacity: Causes, consequences, and a partial solution. *Management Science.*

Hosios, A. J. (1990). On the Efficiency of Matching and Related Models of Search and Unemployment. *Review of Economic Studies*, 57(2):279–298.

Hui, X., Saeedi, M., Shen, Z., and Sundaresan, N. (2016). Reputation and Regulations: Evidence from eBay. *Management Science*, 62(12):3604–3616.

Jolivet, G., Jullien, B., and Postel-Vinay, F. (2016). Reputation and prices on the e-market: Evidence from a major French platform. *International Journal of Industrial Organization*, 45:59–75.

Klein, T. J., Lambertz, C., and Stahl, K. O. (2016). Market Transparency, Adverse Selection, and Moral Hazard. *Journal of Political Economy*, 124(6):1677–1713.

Kranton, R. E. (2003). Competition and the incentive to produce high quality. *Economica*, 70(279):385–404.

Lane, J., Woodworth, R. M., et al. (2016). The sharing economy checks in: An analysis of airbnb in the united states. *CBRE Hotel's Americas Research.*

Peters, M. (1991). Ex ante price offers in matching games non-steady states. *Econometrica: Journal of the Econometric Society*, pages 1425–1454.

Proserpio, D., Xu, W., and Zervas, G. (2018). You get what you give: theory and evidence of reciprocity in the sharing economy. *Quantitative Marketing and Economics*, 16(4):371–407.

Shi, S. and Delacroix, A. (2018). Should buyers or sellers organize trade in a frictional market? *The Quarterly Journal of Economics*, 133(4):2171–2214.

Tadelis, S. (2016). Reputation and feedback systems in online platform markets. *Annual Review of Economics*, 8:321–340.

Wright, R., Kircher, P., Julien, B., and Guerrieri, V. (2017). Directed search: A guided tour. Technical report, National Bureau of Economic Research.

Zervas, G., Proserpio, D., and Byers, J. W. (2017). The rise of the sharing economy: Estimating the impact of airbnb on the hotel industry. *Journal of Marketing Research*, 54(5):687–705.

Appendix

A Model

Here I provide proofs of the Propositions and Theorems discussed in Section 2. Before doing that, I briefly discuss the hosts optimal pricing if the cost of effort becomes public information after being drawn by hosts. I show that this allocation may not be sustained when the cost of effort is hosts' private information. Then, in the context of asymmetry of information, I characterize non-reputational equilibria with separating strategies in prices pointing out the additional conditions that are necessary for their existence. Finally, I characterize the constrained efficient reputational equilibrium allocation and I proceed with the proofs.

A.1 Perfect Information

With public information about hosts' cost of effort, the effort exerted in period 1 and the price posted in period 2 do not impact guests' beliefs. In period 2, the problem for hosts who draw $c = 0$ is defined as follows:

$$\begin{aligned} &\max_{p_2} \; p_2 \alpha(\theta_2) \\ &s.t. \; (a + b - p_2)\frac{\alpha(\theta_2)}{\theta_2} = U_2, \end{aligned}$$

where U_2 is the guests' expected utility for a match with hosts in period 2. Accordingly, the optimal price and tightness for hosts with $c = 0$, p_2^0, θ_2^0 are defined in terms of U_2. If $a + b < U_2$, then $\theta_2^0 = 0$ and $p_2^0 = 0$. If $a + b \geq U_2$:

$$\begin{aligned} \alpha'(\theta_2^0) &= \frac{U_2}{a+b} \\ p_2^0 &= a + b - \frac{\theta_2^0}{\alpha(\theta_2^0)} U_2. \end{aligned} \tag{A.1}$$

Thus, the expected profit with public information for hosts with $c = 0$ is defined as follows:

$$\Pi_2(a + b) = (a + b)(\alpha(\theta_2^0) - \alpha'(\theta_2^0)\theta_2^0), \tag{A.2}$$

where θ_2^0 is defined by Equation A.1. The expected profit is increasing in the guests' surplus of transactions $(a+b)$ if $a+b \geq U_2$:

$$\begin{aligned}
\frac{\partial \Pi_2(a+b)}{\partial (a+b)} &= \alpha(\theta_2^0) - \alpha'(\theta_2^0)\theta_2^0 + (a+b)\frac{\partial(\alpha(\theta_2^0) - \alpha'(\theta_2^0)\theta_2^0)}{\partial \theta_2^0}\frac{\partial \theta_2^0}{\partial (a+b)} \\
&= \alpha(\theta_2^0) - \alpha'(\theta_2^0)\theta_2^0 - (a+b)\alpha''(\theta_2^0)\theta_2^0\frac{\partial \theta_2^0}{\partial (a+b)} \\
&= \alpha(\theta_2^0) - \alpha'(\theta_2^0)\theta_2^0 + (a+b)\alpha''(\theta_2^0)\theta_2^0\frac{1}{\alpha''(\theta_2^0)}\frac{U_2}{(a+b)^2} \\
&= \alpha(\theta_2^0) > 0,
\end{aligned}$$

where the third passage directly follows from the properties of the derivative of the inverse function.[30] Conversely, the expected profit is decreasing in U_2 if $a+b \geq U_2$:

$$\begin{aligned}
\frac{\partial \Pi_2(a+b)}{\partial U_2} &= (a+b)\frac{\partial(\alpha(\theta_2^0) - \alpha'(\theta_2^0)\theta_2^0)}{\partial \theta_2^0}\frac{\partial \theta_2^0}{\partial U_2} \\
&= -(a+b)\alpha''(\theta_2^0)\theta_2^0\frac{\partial \theta_2^0}{\partial U_2} \\
&= -(a+b)\alpha''(\theta_2^0)\theta_2^0\frac{1}{\alpha''(\theta_2^0)}\frac{1}{(a+b)} \\
&= -\theta_2^0 < 0.
\end{aligned}$$

Similarly, in period 2, the problem for hosts who draw $c = k > 0$ is defined as follows:

$$\begin{aligned}
&\max_{p_2} \; p_2\alpha(\theta_2) \\
&s.t. \;\; (b - p_2)\frac{\alpha(\theta_2)}{\theta_2} = U_2.
\end{aligned} \tag{A.3}$$

If $b < U_2$, then the optimal price and tightness for hosts with cost of effort $c = k > 0$ p_2^k, θ_2^k are $p_2^k = 0$ and $\theta_2^k = 0$. If $b \geq U_2$:

$$\begin{aligned}
\alpha'(\theta_2^k) &= \frac{U_2}{b} \\
p_2^k &= b - \frac{\theta_2^k}{\alpha(\theta_2^k)}U_2.
\end{aligned} \tag{A.4}$$

Thus, the expected profit with public information for hosts with $c = k > 0$ is defined as follows:

$$\Pi_2(b) = (b)(\alpha(\theta_2^k) - \alpha'(\theta_2^k)\theta_2^k), \tag{A.5}$$

[30] Recall that the first derivative of the function α is invertible by Assumption 1.

where θ_2^k is defined by Equation A.4. Similarly to the case of hosts with cost of effort $c = 0$, the expected profit is increasing in b and decreasing in U_2 if $b \geq U_2$:

$$\frac{\partial \Pi_2(b)}{\partial b} = \alpha(\theta_2^k) > 0$$
$$\frac{\partial \Pi_2(b)}{\partial U_2} = -\theta_2^k < 0.$$

Hosts who did not match with guests in period 1 do not draw their cost of effort and, together with new arrivals solve the following problem in period 2.

$$\max_{p_2} \; p_2\alpha(\theta_2) \qquad \text{(A.6)}$$
$$s.t. \;\; (a\pi + b - p_2)\frac{\alpha(\theta_2)}{\theta_2} = U_2.$$

The ex-ante guests' utility from a transaction with this class of hosts is $a\pi + b$ since, with probability π hosts draw zero cost of effort and then they exert effort $e_2 = 1$. Otherwise, with probability $1-\pi$ they draw positive cost of effort and they have no incentives to exert effort: $e_2 = 0$. Accordingly, if $a\pi + b < U_2$, the optimal price and tightness for these hosts, $p_2^{\emptyset}$, $\theta_2^{\emptyset}$ are $p_2^{\emptyset} = 0$ and $\theta_2^{\emptyset} = 0$. If $a\pi + b \geq U_2$:

$$\alpha'(\theta_2^{\emptyset}) = \frac{U_2}{a\pi + b} \qquad \text{(A.7)}$$
$$p_2^{\emptyset} = a\pi + b - \frac{\theta_2^{\emptyset}}{\alpha(\theta_2^{\emptyset})}U_2.$$

Thus, the expected profit for hosts who did not match with guests in period 1 and for new arrivals is defined as follows:

$$\Pi_2(a\pi + b) = (a\pi + b)(\alpha(\theta_2^{\emptyset}) - \alpha'(\theta_2^{\emptyset})\theta_2^{\emptyset}), \qquad \text{(A.8)}$$

where $\theta_2^{\emptyset}$ is defined by Equation A.7.

In period 1, hosts have not yet drawn their cost of effort and their problem is the following:

$$\max_{p_1} \; (p_1)\alpha(\theta_1) + \beta\alpha(\theta_1)(\pi\Pi_2(a + b) + (1-\pi)\Pi_2(b)) + \beta(1 - \alpha(\theta_1))\Pi_2(a\pi + b) \qquad \text{(A.9)}$$
$$s.t. \;\; (a\pi + b - p_1)\frac{\alpha(\theta_1)}{\theta_1} = U_1.$$

The ex-ante guests' utility from a transaction in period 1 is $a\pi + b$ since hosts who draw positive cost of effort have no incentives to exert effort: their cost of effort is public information and they cannot commit to exert effort in period 2. Thus, if $a\pi + b + \beta(\pi\Pi_2(a + b) + (1-\pi)\Pi_2(b) - \Pi_2(a\pi + b)) < U_1$, the optimal price and tightness for these hosts, $p_1^\emptyset$, $\theta_1^\emptyset$ are $p_1^\emptyset = 0$ and $\theta_1^\emptyset = 0$. Otherwise:

$$\begin{aligned} \alpha'(\theta_1^\emptyset) &= \frac{U_1}{a\pi + b + \beta(\pi\Pi_2(a+b) + (1-\pi)\Pi_2(b) - \Pi_2(a\pi + b))} \\ p_1^\emptyset &= a\pi + b - \frac{\theta_1^\emptyset}{\alpha(\theta_1^\emptyset)} U_1. \end{aligned} \tag{A.10}$$

If the cost of effort is hosts' private information, the equilibrium above may not be sustained. Hosts are better-off posting p_2^0 relative to p_2^k. This follows since U_2 is the same for hosts with different cost of effort. Thus, $\alpha'(\theta_2^0) < \alpha'(\theta_2^k)$ from Equations A.1 and A.4. Then, by the concavity of α, $\theta_2^0 > \theta_2^k$ and $\alpha(\theta_2^0) > \alpha(\theta_2^k)$: hosts with $c = 0$ have higher chances to be matched with guests relative to hosts with $c = k$. Hence, hosts are better-off posting p_2^0 with expected profits equal to $p_2^0\alpha(\theta_2^0)$:

$$\begin{aligned} p_2^0\alpha(\theta_2^0) &= (a+b)(\alpha(\theta_2^0) - \alpha'(\theta_2^0)\theta_2^0) \\ &> b(\alpha(\theta_2^0) - \alpha'(\theta_2^0)\theta_2^0) > b(\alpha(\theta_2^k) - \alpha'(\theta_2^k)\theta_2^k) \end{aligned}$$

where the inequality in the last passage in due to Assumption 1: $\alpha(\theta) - \alpha'(\theta)\theta > 0 \ \forall\theta$ and $\frac{\partial(\alpha(\theta) - \alpha'(\theta)\theta)}{\partial\theta} = -\alpha''(\theta)\theta > 0$.

Accordingly, the perfect information equilibrium may not be sustained if the cost of effort is hosts' private information. In particular, if the following condition holds, hosts who draw cost $c = k > 0$ are willing to exert effort, incur in cost k and obtain future expected profits $p_2^0\alpha(\theta_2^0)$:

$$\beta(p_2^0\alpha(\theta_2^0) - p_2^k\alpha(\theta_2^k)) \geq c. \tag{A.11}$$

If the condition in A.11 does not hold, then the perfect information equilibrium allocation can be sustained. In the next Section, I will show how this allocation is a particular case of reputational equilibrium.

A.2 Non-Reputational Equilibria

In this paper I focus on reputational equilibria where the information provided by hosts' histories is not made ineffective by the prices posted in period 2. Here I briefly discuss non-reputational equilibria.

As mentioned in the main text, in non-reputational equilibria, hosts with different cost of effort play separate pricing strategies in period 2. Accordingly, guests can perfectly infer hosts' cost of effort observing period 2 prices irrespectively of hosts' histories. In equilibrium, hosts who draw cost $c = k > 0$ post in period 2 the perfect information price p_2^k. Differently, hosts who draw cost $c = 0$ post in period 2 price $p_2^{sep} > 0$ such that hosts with cost $c = k > 0$ are better-off posting p_2^k. In this sense, the existence of non-reputational equilibria relies on the fact that the profit $p_2^k \alpha(\theta_2^k)$ is strictly positive, that is $b > U_2$. If this condition holds, then the following two incentive compatibility constraints have to be satisfied:

$$
\begin{aligned}
p_2^k \alpha(\theta_2^k) &\geq p_2^{sep} \alpha(\theta_2^{sep}) \\
p_2^{sep} \alpha(\theta_2^{sep}) &\geq p_2^k \alpha(\theta_2^k).
\end{aligned}
$$

The ex-ante utility for guests who are matched with hosts posting p_2^k is b; whereas, the utility for those matched with hosts posting p_2^{sep} is $a + b$:

$$
(b - p_2^k)\frac{\alpha(\theta_2^k)}{\theta_2^k} = U_2 = (a + b - p_2^{sep})\frac{\alpha(\theta_2^{sep})}{\theta_2^{sep}}. \tag{A.12}
$$

From the incentive compatibility constraints it results that, when host with $c = 0$ separate, they do not increase their expected profits since $p_2^k \alpha(\theta_2^k) = p_2^{sep} \alpha(\theta_2^{sep})$. In particular, from Equation A.12, $p_2^{sep} > p_2^k$ and $\theta_2^{sep} > \theta_2^k$. Accordingly, the existence of this equilibrium relies on hosts' willingness to separate even when their expected profits do not increase after separating.

A.3 Reputational Equilibria

Here I provide the proofs of the Propositions and Theorems discussed in Section 2. At the same time, I illustrate the constrained efficient allocation and I show that the prices posted by hosts in the equilibrium respect the Hosios (1990) conditions.

The constrained efficient allocation is the allocation that a benevolent social planner would choose taking as given the following elements:

- the frictions that characterize the matching between hosts and guests;
- the hosts' entry cost f;
- the hosts' private information concerning the cost of effort.

Accordingly, the social planner aims at allocating guests to hosts in order to implement the efficient hosts' entry and effort provision.

In line with the main text, I start my analysis from period 2.

A.3.1 Period 2

In order to implement the efficient hosts' entry in period 2, the social planner faces the following problem:

$$\max_{\theta_2}(a\pi + b)\frac{\alpha(\theta_2)}{\theta_2} - \frac{f}{\theta_2}$$

The factor $(a\pi + b)\frac{\alpha(\theta_2)}{\theta_2}$ represents the expected surplus from a transaction for each guest, whereas $\frac{f}{\theta_2}$ defines the hosts' entry costs for each guest. The optimal θ_2^* that maximizes the social planner objective function is such that:

$$(a\pi + b)(\alpha(\theta_2^*) - \alpha'(\theta_2^*)\theta_2^*) = f \tag{A.13}$$

It is possible to note that the optimal price posted by hosts who enter the platform in period 2, $p_2^{\emptyset}$ implements the efficient entry condition of period 2 when the hosts' free entry condition is binding. The optimal expected profits for new entrant hosts is defined by Equation A.8 and it equals the LHS of Equation A.13.

Accordingly, the latter condition equalizes the optimal expected profits for new entrant hosts to the entry costs f: i.e. it imposes a binding free entry condition for entrant hosts in period 2. The rule proposed by Hosios (1990) states that hosts' entry is constrained efficient when the two sides of the market share the ex-ante surplus of transactions (in this case $a\pi + b$) according to the elasticity of the matching function with respect to the tightness. In fact, in this case the expected profits for new entrant hosts, $\Pi_2(a\pi + b)$, and the guests' expected utility, U_2, are defined as follows:

$$\Pi_2(a\pi + b) = p_2^{\emptyset}\alpha(\theta_2^{\emptyset}) = (a\pi + b)(1 - \epsilon_2^{\emptyset})\alpha(\theta_2^{\emptyset})$$
$$U_2 = (a\pi + b)\epsilon_2^{\emptyset}\frac{\alpha(\theta_2^{\emptyset})}{\theta_2^{\emptyset}},$$

where $\epsilon_2^{\emptyset} = \alpha'(\theta_2^{\emptyset})\frac{\theta_2^{\emptyset}}{\alpha(\theta_2^{\emptyset})}$ denotes the elasticity of the matching function with respect to the tightness calculated at $\theta_2^{\emptyset}$.

I analyze the constrained efficient allocation for hosts who enter in period 1 in the next Section. In period 2 they post prices to maximize their profits given guests' beliefs $\bar{\mu}_2(h)$ and U_2.

Proof of Proposition 1. Assuming that guests do not update their beliefs about hosts' cost of effort after observing prices in period 2, hosts who were matched in period 1 and with history h solve the following problem in period 2:

$$\max_{p_2} \; p_2\alpha(\theta_2) \tag{A.14}$$
$$s.t. \; (a\bar{\mu}_2(h) + b - p_2)\frac{\alpha(\theta_2)}{\theta_2} = U_2.$$

If $a\bar{\mu}_2(h) + b < U_2$, then the optimal price and tightness $p_2^{pool}(h)$, $\theta_2^{pool}(h)$ are $p_2^{pool}(h) = 0$ and $\theta_2^{pool}(h) = 0$. If $a\bar{\mu}_2(h) + b \geq U_2$:

$$\alpha'(\theta_2^{pool}) = \frac{U_2}{a\bar{\mu}_2(h) + b} \tag{A.15}$$
$$p_2^{pool} = a\bar{\mu}_2(h) + b - \frac{\theta_2^{pool}}{\alpha(\theta_2^{pool}}U_2.$$

Similarly, hosts who were not matched in period 1 and new entrants solve the problem in Equation A.6 and their optimal price and tightness $p_2^{\emptyset}(h), \theta_2^{\emptyset}(h)$ are reported in Equation A.7.

□

It is possible to note that the optimal prices for hosts who enter in period 1 follow Hosios (1990) conditions since hosts and guests share the ex-ante surplus $a\bar{\mu}_2(h) + b$ according to the elasticity of the matching function. In particular:

$$\Pi_2(a\bar{\mu}_2(h) + b) = p_2^{pool}(h)\alpha(\theta_2^{pool}(h)) = (a\bar{\mu}_2(h) + b)(1 - \epsilon_2^{pool}(h))\alpha(\theta_2^{pool}(h))$$
$$U_2 = (a\bar{\mu}_2(h) + b)\epsilon_2^{pool}(h)\frac{\alpha(\theta_2^{pool}(h))}{\theta_2^{pool}(h)},$$

where $\epsilon_2^{pool}(h) = \alpha'(\theta_2^{pool}(h))\frac{\theta_2^{pool}(h)}{\alpha(\theta_2^{pool}(h))}$ denotes the elasticity of the matching function with respect to the tightness calculated at θ_2^{pool}.Accordingly, the ex-ante surplus of transactions in period 2 is greater for hosts who exert effort in period 1. Furthermore, hosts who exert effort in period 1 get a greater share of the surplus since the elasticity $\epsilon_2^{pool}(h)$ is decreasing in the tightness and $\theta_2^{pool}(h) > \theta_2^k \; \forall\bar{\mu}_2(h)0$. In this sense, in order to increase the effort provision (in period 1) and obtain the efficient hosts' entry in period 2, the social planner may commit to allocate guests to hosts in period 2 such that the tightness levels $\theta_2^{pool}(h), \theta_2^{\emptyset}$ 2 are formed.

A.3.2 Period 1

In line with the analysis in the main text, I start with the proof of Proposition 2 regarding the effort provision in period 1. Then, I provide the proof for Proposition 3 and I characterize the constrained efficient allocation in period 1.

Proof of Proposition 2. The effort strategy by hosts with $c = k > 0$ in period 1 realizes the interest of these hosts to mimic hosts with $c = 0$, post higher prices and attract more guests as it has been observed in Proposition 1. In particular, exerting $e_1 = 1$, hosts with $c = k > 0$ pool together with hosts with $c = 0$ in period 2, posting $p_2^{pool}(e_1 = 1)$. Exerting $e = 0$, with $c = k > 0$ cannot pool anymore since their history is fully revealing their costs. The value of pooling depends on the guests' interim beliefs. They are derived by the Bayes formula when possible:

$$\bar{\mu}_2(e_1 = 1) = \frac{\pi}{\pi + (1-\pi)\omega} \tag{A.16}$$

$$\bar{\mu}_2(\emptyset) = \pi \tag{A.17}$$

$$\bar{\mu}_2(e_1 = 0) = 0, \tag{A.18}$$

where $\omega \in [0, 1]$ is the probability to exert effort $e_1 = 1$ by hosts with $c = k > 0$ in equilibrium in period 1. In this sense, the discounted marginal benefits of exerting effort for non-commitment types are defined as follows:

$$MB = \beta(p_2^{pool}(e_1 = 1)\alpha(\theta_2^{pool}(e_1 = 1)) - p_2^{pool}(e_1 = 0)\alpha(\theta_2^{pool})(e_1 = 0)).$$

Recalling the function $\Pi_2(.)$ introduced in the previous Section, the discounted marginal benefits can be defined as follows:

$$MB = \beta\Big(\Pi_2\big(a\frac{\pi}{\pi + (1-\pi)\omega} + b\big) - \Pi_2(b)\Big),$$

where the function $\Pi_2(.)$ is weakly increasing in the value of $a\bar{\mu}_2(h) + b$. Hosts with $c = k > 0$ compare MB with the cost of effort k. The following algorithm characterizes the equilibrium level of ω:

1. Consider the case $\omega = 1$ and calculate the MB . If MB is greater than k:

   $$\beta\Big(\Pi_2\big(a\frac{\pi}{\pi + (1-\pi)\omega} + b\big) - \Pi_2(b)\Big) \geq k,$$

   then, in equilibrium hosts with $c = k > 0$ exert effort in period 1 with probability $\omega = 1$;

2. If the inequality above does not hold true, then consider the case $\omega = 0$ and calculate again MB. If MB is lower than k:

   $$\beta\Big(\Pi_2\big(a\frac{\pi}{\pi + (1-\pi)\omega} + b\big) - \Pi_2(b)\Big) \leq k,$$

   then, in equilibrium hosts with $c = k > 0$ exert effort in period 1 with probability $\omega = 0$;

3. If the two inequalities above do not hold true, then derive $\omega \in (0,1)$ such that the following equality holds:

$$\beta\Big(\Pi_2\big(a\frac{\pi}{\pi+(1-\pi)\omega}+b\big)-\Pi_2(b)\Big)=k, \tag{A.19}$$

Since the *LHS* of Equation A.19 is strictly decreasing in ω, it admits only one solution.

□

Proof of Proposition 3. Hosts who enter in period 1 have not yet drawn their cost of effort. Thus, their problem in period 1 is the following:

$$\begin{aligned}
\max_{p_1}\ & (p_1-k(1-\pi)\omega)\alpha(\theta_1) \\
& +\beta\alpha(\theta_1)\big(\pi\Pi_2\big(a\frac{\pi}{\pi+(1-\pi)\omega}+b\big)+(1-\pi)(1-\omega)\Pi_2(b)\big) \\
& +\beta(1-\alpha(\theta_1))\Pi_2(a\pi+b) \\
s.t.\ & (a(\pi+(1-\pi)\omega)+b-p_1)\frac{\alpha(\theta_1)}{\theta_1}=U_1.
\end{aligned}$$

The ex-ante guests' utility from a transaction in period 1 is $a(\pi+(1-\pi)\omega)+b$ since hosts who draw positive cost of effort exert effort in period 1 with probability ω. Thus, if $a(\pi+(1-\pi)\omega)+b-k(1-\pi)\omega+\beta\Delta\Pi<U_1$, the optimal price and tightness for these hosts, $p_1^{\emptyset},\theta_1^{\emptyset}$ are $p_1^{\emptyset}=0$ and $\theta_1^{\emptyset}=0$. Otherwise:

$$\begin{aligned}
\alpha'(\theta_1^{\emptyset}) &= \frac{U_1}{a(\pi+(1-\pi)\omega)+b-k(1-\pi)\omega+\beta\Delta\Pi} \\
p_1^{\emptyset} &= a(\pi+(1-\pi)\omega)+b-\frac{\theta_1^{\emptyset}}{\alpha(\theta_1^{\emptyset})}U_1.
\end{aligned}$$

In this sense, if $a(\pi+(1-\pi)\omega)+b-k(1-\pi)\omega+\beta\Delta\Pi\geq U_1$, the expected profits for new entrants in period 1 are defined as follows:

$$\Pi_1=(a(\pi+(1-\pi)\omega)+b-k(1-\pi)\omega+\beta\Delta\Pi)(\alpha(\theta_1^{\emptyset})-\alpha'(\theta_1^{\emptyset})\theta_1^{\emptyset})+\beta\Pi_2(a\pi+b). \tag{A.20}$$

□

The constrained efficient allocation in period 1 implies the efficient hosts' entry and effort provision in period 1. Accordingly, the social planner commits to allocate guests to hosts in period 2 in order to form the tightness levels $\theta_2^{pool}(h),\theta_2^{\emptyset}$. In this sense, hosts who draw cost $c=k>0$ in period 1 have incentives to exert effort with probability ω in line with Proposition 2. Therefore, the social planner solves the following problem in period 1:

$$\max_{\theta_1}(a(\pi+(1-\pi)\omega)+b-k(1-\pi)\omega)\frac{\alpha(\theta_1)}{\theta_1}+\frac{R}{\theta_1}-\frac{f}{\theta_1}$$

Similarly to period 2, the factor $a(\pi+(1-\pi)\omega)+b-k(1-\pi)\omega)\frac{\alpha(\theta_1)}{\theta_1}$ represents the expected surplus from a transaction for each guest, whereas $\frac{f}{\theta_1}$ defines the hosts' entry costs for each guest. Yet, in period 1, an additional element forms the surplus of a transaction. The factor R captures the value of a transaction in updating hosts' reputation and changing the ex-ante surplus of transactions in period 2:

$$R=\beta\alpha(\theta_1)(\pi\Pi_2\big(a\frac{\pi}{\pi+(1-\pi)\omega}+b\big)+(1-\pi)(1-\omega)\Pi_2(b)\big)+\beta(1-\alpha(\theta_1))Pi_2(a\pi+b).$$

The optimal θ_1^* that maximizes the social planner objective function is such that:

$$(a(\pi+(1-\pi)\omega)+b-k(1-\pi)\omega+\beta\Delta\Pi)(\alpha(\theta_1^*)-\alpha'(\theta_1^*)\theta_1^*)+\beta\Pi_2(a\pi+b)=f. \tag{A.21}$$

It is possible to note that the optimal price posted by hosts who enter the platform in period 1, $p_1^\emptyset$ implements the efficient entry condition of period 1 when the hosts' free entry condition is binding. The optimal expected profits for new entrant hosts is defined by Equation A.20 and it equals the LHS of Equation A.20. Accordingly, the latter condition equalizes the optimal expected profits for new entrant hosts to the entry costs f: i.e. it imposes a binding free entry condition for entrant hosts in period 1.

A.3.3 Existence and Uniqueness

The proof of Theorem 1 has the following structure: first, I assume that a positive measure of hosts enter in period 2. With this assumption, I show the existence and the uniqueness of the equilibrium and I derive the threshold level $\bar{G}$ such that there is entry in the second period for $G>\bar{G}$.

Proof of Theorem 1. With a positive measure of hosts entering in period 2, the free entry condition for hosts in period 2 holds with equality. From the free entry condition and the pricing problem for new entrants (Proposition 1), it is possible to uniquely determine $p_2^\emptyset, \theta_2^\emptyset$ and U_2. Accordingly, the free entry condition can be written in terms of $\theta_2^\emptyset$:

$$(a\pi+b)(\alpha(\theta_2^\emptyset)-\theta_2^\emptyset\alpha'(\theta_2^\emptyset))=f. \tag{A.22}$$

From Equation A.22, the equilibrium value of $\theta_2^\emptyset$ can be uniquely derived. Recall that $a\pi+b>0$, $\alpha''(\theta)<0$, and $\alpha(\theta)-\theta\alpha'(\theta)>0\ \forall\theta$. Moreover, $\alpha(\theta)-\theta\alpha'(\theta)$ is strictly increasing in θ. Then, the LHS of Equation A.22 is strictly increasing in θ. For $\theta = 0$, LHS is zero, and the equilibrium value of $\theta_2^{\emptyset}$ is unique and strictly positive with $f > 0$. Using $\theta_2^{\emptyset}$, the equilibrium values of $p_2^{\emptyset}$ and U_2 can be uniquely derived from Equation A.7. With U_2, the values of $p_2^{pool}(h^0)$ and $\theta_2^{pool}(h^0)$ can be uniquely derived by Equation A.4:

$$\alpha'(\theta_2^{pool}(h^0)) = \alpha'(\theta_2^k) = \frac{U_2}{b}$$
$$p_2^{pool}(h^0) = p_2^k = b - \frac{\theta_2^k}{\alpha(\theta_2^k)} U_2,$$

if $b \geq U_2$. Otherwise $\theta_2^{pool}(h^0) = 0$ and $p_2^{pool}(h^0) = 0$. Similarly, the values of $p_2^{pool}(h^1)$ and $\theta_2^{pool}(h^1)$ can be uniquely derived in terms of ω:

$$\alpha'(\theta_2^{pool}(h^1)) = \frac{U_2}{a\frac{\pi}{\pi+(1-\pi)\omega} + b}.$$
$$p_2^{pool}(h^1) = a\frac{\pi}{\pi + (1-\pi)\omega} + b - \frac{\theta_2^{pool}(h^1)}{\alpha(\theta_2^{pool}(h^1))} U_2.$$

As showed early, $\theta_2^{pool}(h^{\emptyset}) > 0$, and $a\pi + b - c > U_2$. Still, since $\frac{\pi}{\pi+(1-\pi)\omega} \geq \pi$, then we have that $a\frac{\pi}{\pi+(1-\pi)\omega} + b > U_2$ and $\theta_2^{pool}(h^1) > 0$. By Proposition 2, ω can be uniquely determined. It follows that also $\theta_2^{pool}(h^1)$ and $p_2^{pool}(h^1)$ are uniquely determined. Accordingly, the equilibrium system of equations uniquely determines all terms regarding period 2.

The expected profits for entrant hosts in period 1 can be rewritten as follows:

$$(p_1^{\emptyset} - k(1-\pi)\omega)\alpha(\theta_1^{\emptyset}) + \beta\alpha(\theta_1^{\emptyset})\Delta\Pi + \beta\Pi_2(a\pi + b),$$

where $\Delta\Pi$ is defined in Proposition 3 and denotes the value of a transaction in terms of reputation updating. Then, by Proposition 3, with $\theta_1^{\emptyset} > 0$:

$$(p_1^{\emptyset} - k(1-\pi)\omega)\alpha(\theta_1^{\emptyset}) + \beta\alpha(\theta_1^{\emptyset})\Delta\Pi = [a(\pi(1-\pi)\omega) + b + \beta\Delta\Pi - k\omega(1-\pi)](1 - \epsilon_1^{\emptyset}),$$

where $\epsilon_1^{\emptyset} = \alpha'(\theta_1^{\emptyset})\frac{\theta_1^{\emptyset}}{\alpha(\theta_1^{\emptyset})}$ denotes the elasticity of the matching function with respect to the tightness calculated at $\theta_1^{\emptyset}$. Thus, the free entry condition in period 1 has the following structure:

$$[a(\pi(1-\pi)\omega) + b + \beta\Delta\Pi - k\omega(1-\pi)](1 - \epsilon_1^{\emptyset}) + \beta\Pi_2(a\pi + b) = f. \tag{A.23}$$

Then, the optimal value of $\epsilon_1^{\emptyset}$ and $\theta_1^{\emptyset}$ can be uniquely derived by Equation A.23. It is possible to note that: $a(\pi(1-\pi)\omega) + b + \beta\Delta\Pi - k\omega(1-\pi) \geq 0$ for all values of θ_1; the value of ϵ_1 is strictly decreasing in θ_1 and $\Pi_2(a\pi + b) \leq f$ by the free-entry condition in period 2. Accordingly, Equation A.23 uniquely characterizes $\theta_1^{\emptyset}$ with $\theta_1^{\emptyset} > 0$. Knowing $\theta_1^{\emptyset}$, I obtain $p_1^{\emptyset}$ and U_1 by Proposition 3, and the measure of entrants in period 1, n_1, by $\theta_1^{\emptyset} = \frac{1}{n_1}$. With n_1, ω, and θ_1^{pool} the measures of hosts who entered in period 1 and have with histories h^1, h^0, and $h^{\emptyset}$ in period 2 are derived as follows:

$$\begin{aligned} n_2(h^1) &= (\omega n_1(1-\pi) + \pi n_1)\alpha(\theta_1^{\emptyset}) \\ n_2(h^0) &= (1-\omega)n_1(1-\pi)\alpha(\theta_1^{\emptyset}) \\ n_2(h^{\emptyset}) &= n_1(1-\alpha(\theta_1^{\emptyset})). \end{aligned}$$

Then, with $\theta_2^{pool}(h^1), \theta_2^{pool}(h^0)$, and $\theta_2^{\emptyset}$, the measures of guests who direct their search to hosts with histories h^1, h^0, and $h^{\emptyset}$ posting $p_2^{pool}(h^1), p_2^{pool}(h^0)$ and $p_2^{\emptyset}$, are respectively the following:

$$\begin{aligned} g_2(h^1) &= \theta_2^{pool}(h^1)n_2(h^1) \\ g_2(h^0) &= \theta_2^{pool}(h^0)n_2(h^0) \\ g_2(h^{\emptyset}) &= G - g_2(h^1) - g_2(h^0). \end{aligned}$$

Finally, the number of new entrants in period 2 is the difference between the total measure of hosts with history $h^{\emptyset}$ and $n_2(h^{\emptyset})$:

$$\frac{g_2(h^{\emptyset})}{\theta_2^{\emptyset}} - n_2(h^{\emptyset}). \tag{A.24}$$

I started the proof assuming that a positive measure of hosts enter in period 2. Still, for some G, the value in Equation A.24 can be negative. In this sense, the proof of the existence and the uniqueness of the equilibrium relies on a value of $G \geq \bar{G}$, with $\bar{G}$:

$$\bar{G} = g_2(h^1) + g_2(h^0) + n_2(h^{\emptyset})\theta_2^{\emptyset}. \tag{A.25}$$

□

A.3.4 Testable Predictions

Proof of Proposition 4. The measure of guests present in the market in period 2 is assumed to be big enough to allow hosts' entry in period 2 for both equilibria. Accordingly, the free entry condition is binding for f and f'. Then, $\theta_2^{'\emptyset} > \theta_2^{\emptyset}$ recalling that the expected profits for new entrants is strictly increasing in θ. Moreover, directly from the relationship established in the Proposition 1 between the tightness $\theta_2^{\emptyset}$ and the level of U_2, it results that $U_2 > U_2'$. Accordingly, $\theta_2^{'pool}(h^1) > \theta_2^{pool}(h^1)$ and $\theta_2^{'pool}(h^0) > \theta_2^{pool}(h^0)$ from the Equations in Proposition 1. Accordingly, higher entry costs reduce the number of hosts who enter the market in period 2; thus, the tightness for all hosts increases and the guests' expected utility from the matches decreases. The derivative of the profits over U_2

has already been discussed in Section A.1 using the definition of the function $\Pi_2(.)$. In particular:

$$\begin{aligned}\frac{\partial \Pi_2(a\bar{\mu}_2(h)+b)}{\partial U_2} &= \frac{\partial \Pi_2(a\bar{\mu}_2(h)+b)}{\partial \theta_2^{pool}(h)}\frac{\partial \theta_2^{pool}(h)}{\partial U_2} \\ &= -\theta_2^{pool}(h)\alpha''(\theta_2^{pool}(h))\left[\alpha'^{-1}\left(\frac{U_2}{a\bar{\mu}_2(h)+b}\right)\right]' = -\theta_2^{pool}(h),\end{aligned}$$

if $U_2 \leq a\bar{\mu}_2 + b$. The last passage directly follows from the properties of the derivative of the inverse function. If $U_2 > a\bar{\mu}_2 + b$, the derivative is equal to zero. Similarly, for those hosts who do not have a transaction in period 1:

$$\frac{\partial \Pi_2(a\pi)+b)}{\partial U_2} = -\theta_2^{\emptyset},$$

if $U_2 \leq a\pi + b$. Otherwise, the derivative is equal to zero. Accordingly, a decrease in U_2 has a greater, positive impact on the expected profits for those hosts with a higher value of θ_2. Taking advantage of this result, it is possible to show that $\omega' \geq \omega$ with the same algorithm used in the proof of Proposition 2:

1. Consider the case in which hosts with $c = k > 0$ exert effort with probability $\omega = 1$ when entry costs are f. Then, in equilibrium:

$$\beta(\Pi_2(a\pi + b) - \Pi_2(b)) \geq k. \tag{A.26}$$

   From the previous results about the derivative of profits in period 2, the LHS of Equation A.26 is greater with f'. Then, hosts with $c = k > 0$ exert effort with probability 1 also with entry costs f: $\omega' = 1$;

2. Consider the case $\omega = 0$ when entry costs are f. Then, in equilibrium:

$$\beta(\Pi_2(a + b) - \Pi_2(b) \leq k. \tag{A.27}$$

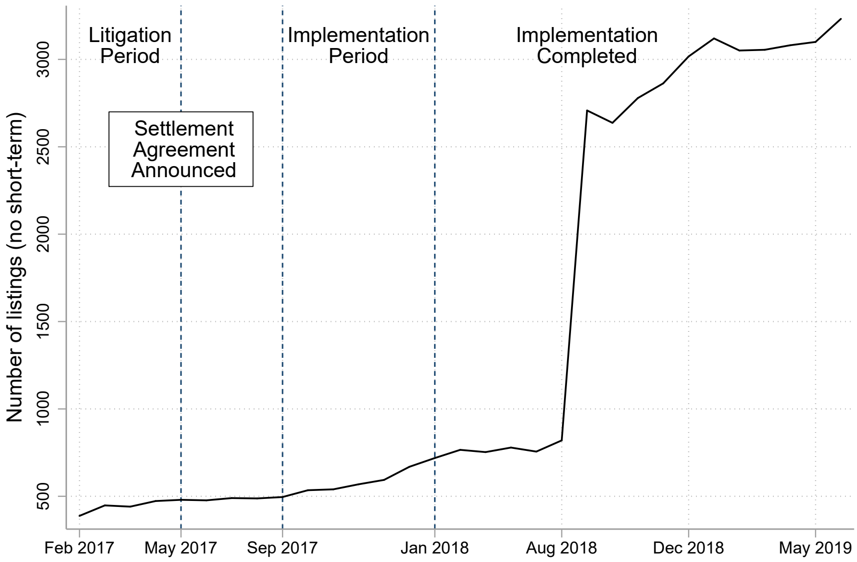
   As before, the LHS of Equation A.27 is greater with f'. Then, Equation A.27 may not be satisfied with f' and, in equilibrium $\omega' \geq 0$;

3. Finally, consider the case in which $\omega \in (0, 1)$ when entry costs are f, such that Equation A.19 is satisfied. With f' the LHS of Equation A.19 increases if $\omega' = \omega$. To restore the equality, the value of ω' has to increase (if possible) so as to decrease the reputation of hosts with history h^1. Thus $\omega' \geq \omega$.

□

B Empirical Setting and Dataset

Figure A.1: Long-Term Airbnb Listings over Time



Note: The figure plots the total number of Airbnb listings that do *not* offer short-term lodging (long-term) in San Francisco over time (different snapshots) from February 2017 to June 2019. The three vertical lines regard the timing of the Settlement Agreement between the San Francisco City Council and Airbnb. The agreement was signed in May 2017 and it has been effective since September 2017. According to the resolution, from January 2018 all eligible Airbnb listings should be registered.

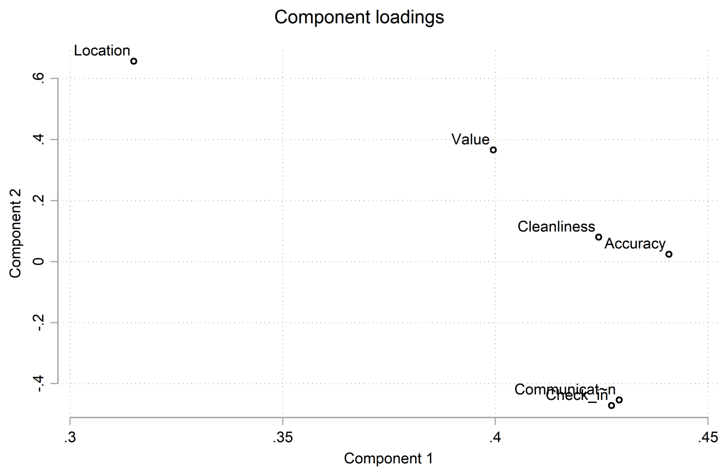
Table A.1: Summary Statistics: the Settlement Agreement and Listings Selection in September 2017.

	Group C		Group D			
	Mean	SD	Mean	SD	Δ	$p-value$
Days in Airbnb	1,054.5	388.5	570.9	242.8	483.5	0.0
Days in Airbnb before September 2017	524.4	286.2	493.4	247.9	30.3	0.0
Total number of reviews	45.5	61.4	9.8	24.4	35.1	0.0
Price per night	206.4	165.0	246.0	232.4	-39.6	0.0
Availability next 30 days	7.1	8.7	3.5	8.4	3.5	0.0
Average rating per snapshot: overall	94.3	6.0	93.1	8.9	1.3	0.0
Average rating per snapshot: accuracy	9.6	0.6	9.4	0.9	0.1	0.0
Average rating per snapshot: check-in	9.8	0.5	9.6	0.6	0.1	0.0
Average rating per snapshot: cleanliness	9.5	0.7	9.2	1.19	0.3	0.0
Average rating per snapshot: communication	9.8	0.5	9.7	0.80	0.9	0.0
Average rating per snapshot: location	9.5	0.6	9.4	0.9	0.08	0.00
Average rating per snapshot: value-for-money	9.2	0.7	9.1	1.5	0.9	0.00
Minimum nights required	5.5	10.0	3.0	4.6	2.4	0.0
No short-term rentals	10%	-	1%	-	0.08	-
Registration displayed or not required	20%	-	3%	-	0.16	-
Number of listings	4,560	-	3,642	-	-	-

Note: The table compare the profile of listings before and after the Settlement Agreement. All the statistics refer to the snapshot regarding September 2017. Listings are divided in two groups: Group C contains all listings who enter the platform before September 2017 and exit after January 2018, when the implementation of the Settlement Agreement was completed. Group D contains all listings who enter the platform before September 2017 and exit before January 2018. The last two columns provide the differences between the statistics' averages and the $p-value$ of the difference. The numbers of listings in the two groups are not equal to the ones shown in Table 2 since not all listings in the two groups were active (present on the platform) at the date of the snapshot regarding September 2017.

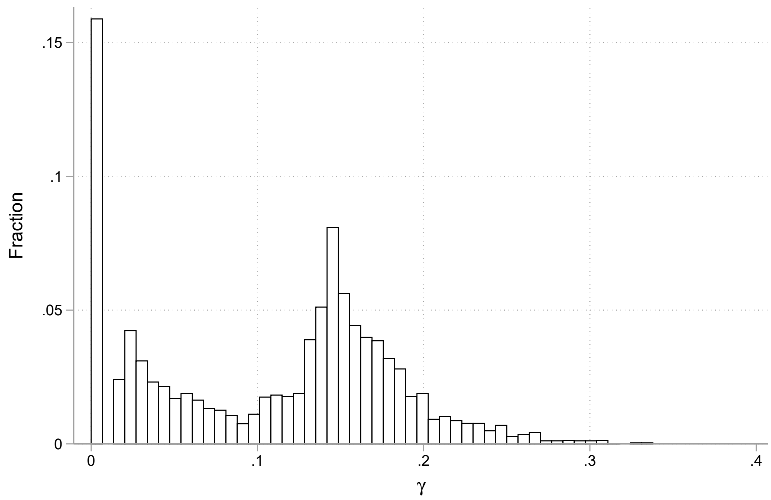
C Identification Strategy

Figure A.2: (PCA) Loading of All Rating Categories over the First Two Components



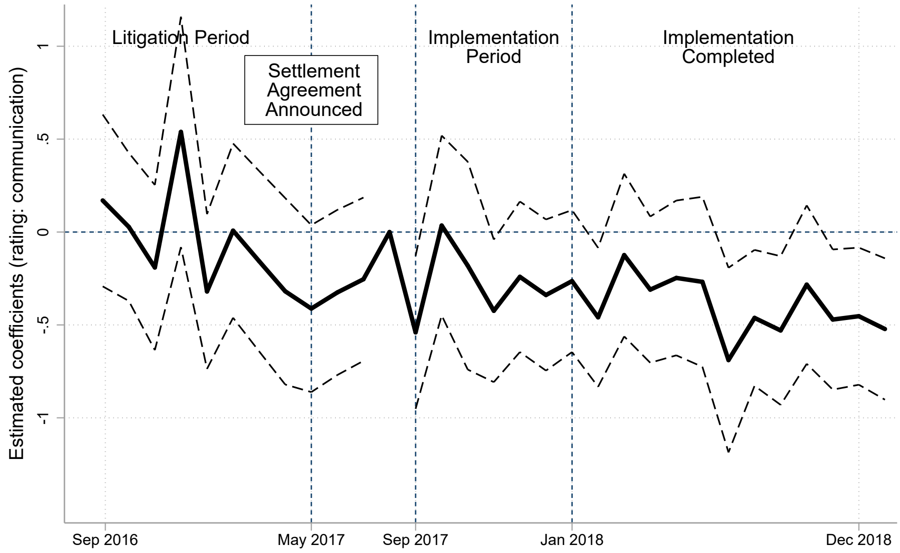
Note: The figure plots the loading of all categories over the first two components of a PCA performed over the rating per snapshot of all the categories. Only listings offering short-term lodging that enter the platform before September 2017 and exit after January 2018 are considered. The ratings are all very correlated and the first component already explain more than 30 percent of the ratings variations. Ratings regarding check-in and communication correlate the most and their loadings are distant from all the others. The rating regarding location moves separately, whereas all the other dimensions tend to have similar loadings. These three results are robust to many specifications of principal components and factor analyses.

Figure A.3: Distribution of γ_i^1



Note: The figure shows the distribution of values of γ_i^1 for the set of listings offering short-term lodging that enter the platform before September 2017 and exit after January 2018 (Group C in Section 3.4). More than 15 percent of listings report a value of $\gamma_i^1 = 0$ with no registered competitors within 1 kilometer in September 2017.

Figure A.4: Estimated Coefficients from Equation 4.4: Ratings Regarding Communication

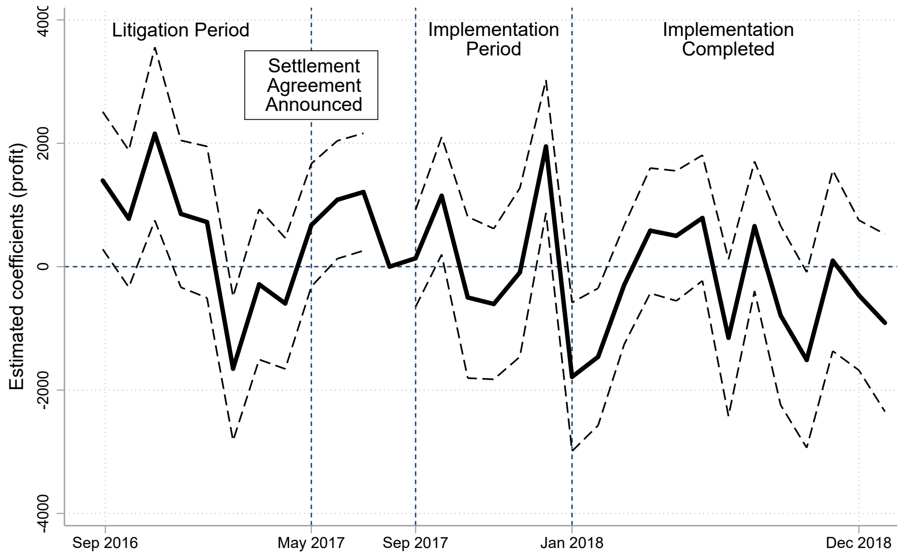


Note: In line with Equation 4.4, $\bar{r}_{i,t}^{comm}$ is regressed on listing and snapshot fixed effects, and on the products between γ_i^j and snapshot dummies. Standard errors are clustered by listing. The graph plots the estimated coefficients on these products. The value of the coefficient corresponding to August 2017 ($\hat{\beta}_{Aug2017}$) is normalized to zero. Only listings offering short-term lodging that enter the platform before September 2017 and exit after January 2018 are considered.

D Extensions

D.1 Competition and Profits

Figure A.5: Estimated Coefficients from Equation 6.1



Note: In line with Equation 6.1, $\pi_{i,t}$ is regressed on listing and snapshot fixed effects, and on the products between γ_i^1 and snapshot dummies. Standard errors are clustered by listing. The graph plots the estimated coefficients on these products. The value of the coefficient corresponding to August 2017 ($\hat{\beta}_{Aug2017}$) is normalized to zero. Only listings offering short-term lodging that enter the platform before September 2017 and exit after January 2018 are considered.

E Robustness Check

E.1 Effort Estimation

The effort estimation presented in Section 6 relies on a relationship between guests' characteristics regarding different features of the same lodging service. In particular, Equation 7.3 assumes a linear model between $guest_{i,t}^{effort}$ and $guest_{i,t}^{location}$ for all Airbnb guests in the dataset (up to the error term $\epsilon_{i,t}$).

Still, the relationship between guests' perception for the components of hosts' services may be heterogeneous for different types of guests; and the assumption of Equation 7.3 may be partially relaxed to account for such heterogeneity. The control function approach derived in Equation 7.4 relies on the assumption that host's effort $e_{i,t}$ can be identified looking at time variations of ratings $\bar{r}_{i,t}^{effort}$ after removing the trend due to the correlation with $\bar{r}_{i,t}^{location}$. Accordingly, this estimation technique cannot allow for time varying parameters affecting the linear relationship between the

Table A.2: OLS Estimates of the Impact of Competition on Hosts' Effort

	$e^{check-in}_{i,t}$			$e^{comm}_{i,t}$		
	(1)	(2)	(3)	(4)	(5)	(6)
$\ln(L^{0.5}_{i,t})\times 10$	-0.0720			-0.0182		
	[0.122]			[0.117]		
$\ln(L^{1}_{i,t})\times 10$		-0.206			-0.136	
		[0.165]			[0.167]	
$\ln(L^{2}_{i,t})\times 10$			-0.321			-0.0748
			[0.222]			[0.230]
Constant	-0.0122	0.0739	0.181	-0.0478	0.0224	-0.00353
	[0.0563]	[0.0959]	[0.156]	[0.0545]	[0.0980]	[0.162]
Listing FE	✓	✓	✓	✓	✓	✓
Snap FE	✓	✓	✓	✓	✓	✓
Mean	0.006	0.006	0.006	0.007	0.007	0.007
R^2	0.00187	0.00136	0.00114	0.00223	0.00170	0.00213
N	55,587	55,587	55,587	55,587	55,587	55,587

Note: Only listings offering short-term lodging that enter the platform before September 2017 and exit after January 2018 are considered. Standard errors clustered by listing are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

two ratings: i.e. α and β cannot vary over time. Still, the guests' perception for the components of hosts' services may differ for specific time-invariant groups of listings. In particular, it is possible to substitute the assumption in Equation 7.3 with the following two-stage formulation in which the parameters α and β differ for each group n:

$$guest^{effort}_{i,t} = \alpha_n + \beta_n guest^{location}_{i,t} + \epsilon_{i,t} \tag{E.1}$$

$$\alpha_n = \alpha + v_n \tag{E.2}$$

$$\beta_n = \beta + u_n, \tag{E.3}$$

where v_n and u_n are the random effect and coefficient varying for each group n. In line with the approach used in Section 6, I derive the following panel regression:

$$\bar{r}^{effort}_{i,t} = \alpha_n - \beta_n\theta_i + \beta_n\bar{r}^{location}_{i,t} + e_{i,t} + \epsilon_{i,t}. \tag{E.4}$$

The main difference between Equation 7.4 and E.4 regards the coefficients α_n and β_n that vary for different groups. Yet, Equation E.4 can be simplified operating a within transformation to account for the listing fixed effect due to the fixed characteristics θ_i:

$$(\bar{r}_{i,t}^{effort} - \bar{\bar{r}}_i^{effort}) = \beta_n(\bar{r}_{i,t}^{location} - \bar{\bar{r}}_i^{location}) + (e_{i,t} - \bar{e}_i) + (\epsilon_{i,t} - \bar{\epsilon}_i) \tag{E.5}$$

$$(\bar{r}_{i,t}^{effort} - \bar{\bar{r}}_i^{effort}) = \beta(\bar{r}_{i,t}^{location} - \bar{\bar{r}}_i^{location}) \tag{E.6}$$

$$+ u_n(\bar{r}_{i,t}^{location} - \bar{\bar{r}}_i^{location}) + (e_{i,t} - \bar{e}_i) + (\epsilon_{i,t} - \bar{\epsilon}_i), \tag{E.7}$$

where $\bar{\bar{r}}_i^{effort}$ and $\bar{\bar{r}}_i^{location}$ are the listing average ratings per snapshot regarding effort and location, respectively.

The within formulation of the panel regression removes the fixed part of the model $\alpha_n - \beta_n\theta_i$ and only the random coefficient u_n remains to capture the heterogeneous relationship between $guest_{i,t}^{effort}$ and $guest_{i,t}^{location}$. Treating u_n as random implies the necessity to have further assumptions about the distribution of the random effect and its independence. In particular, the following assumptions need to hold:

$$E[e_{i,t}u_n|\theta_i] = 0 \tag{OC_3}$$

$$E[\epsilon_{i,t}u_n|\theta_i] = 0 \tag{OC_4}$$

$$E[\bar{r}_{i,t}^{location}u_n|\theta_i] = 0, \tag{OC_5}$$

with $u_n \sim \mathcal{N}(0, \sigma_u^2)$. In this sense, hosts are assumed to not respond to changes in u_n with variations in effort; and u_n is assumed to not be correlated with variations in the rating regarding location.

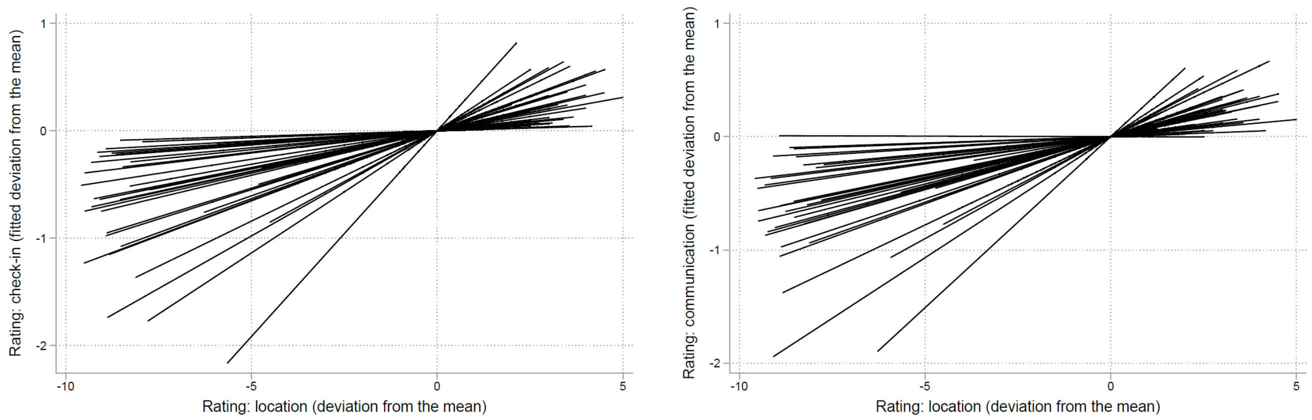
Different time-invariant groups can be used to add heterogeneity in the relationship between guests' characteristics. Here I use the thirty-seven neighborhoods in the San Francisco city center to capture the different profile of guests using Airbnb in the city. In certain neighborhoods, tourists may give extra importance to listings' location; whereas other areas may attract guests with different tastes and priorities. Table A.3 shows the results about the panel fixed effect regression in Equation 7.4 and E.4 for the ratings regarding check-in and communication. The values of β are positive and significant in all cases. This is in line with the assumption of the relationship between the ratings regarding effort and location. Moreover, the Likelihood Ratio test at the bottom of the table shows that the hypothesis of $\sigma_u^2 = 0$ (equivalent to $u_n = 0\ \forall n$) is rejected suggesting that the random slope for each neighborhood improves the predictive power of the model. In line with this result, Figure A.6 shows how the slopes β_n vary for different neighborhoods. In particular, the two graphs plot the fitted values of $(\bar{r}_{i,t}^{effort} - \bar{\bar{r}}_i^{effort})$ over $(\bar{r}_{i,t}^{location} - \bar{\bar{r}}_i^{location})$ using the estimated β and u_n. Other specifications with a wider range of random effects are possible. Still, adding further heterogeneity does not seem to improve the power of the model: when other time-invariant heterogeneity is added (such as the types of rented property) the Likelihood Ratio test shows that the other sources of heterogeneity do not improve the model's predictive power.

Table A.3: Relationship between Guests' Characteristics

	$\bar{r}_{i,t}^{check-in}$		$\bar{r}_{i,t}^{comm}$	
	(1)	(2)	(3)	(4)
$\bar{r}_{i,t}^{location}$	0.0758***	0.0865***	0.0774***	0.0836***
	[0.00219]	[0.0134]	[0.00217]	[0.0113]
u_n		0.0768***		0.0643***
		[0.0108]		[0.00919]
Listing FE	✓	✓	✓	✓
Mean	9.834	9.834	9.846	9.846
LR test vs. linear model		392.89		275.33
N	120,905	120,905	120,905	120,905

Note: Standard errors clustered by listing are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure A.6: Relationship between Guests' Characteristics in Different Neighborhoods



Note: The graphs plots the estimated random coefficients u_n for different neighborhoods. As observed in Table A.3, the relationship between guests' characteristics (effort and location) significantly vary among neighborhoods.

I use the residuals obtained from Equation E.4 as a new measure of host's effort. Repeating the analysis described in Section 6, I get qualitatively similar results. In particular, I regress the variable hosts' response rate over the effort measures $e_{i,t}^{check}$, $e_{i,t}^{comm}$, and the location rating $\bar{r}_{i,t}^{location}$ controlling for listing fixed effects (see Table A.4). As for the previous estimates of the host's effort, the results support condition OC_2: hosts' response rate is positively and significantly correlated with the effort dimensions.

Finally, I replicate again the identification strategy in Section 4 using the new estimates of $e_{i,t}^{check}$ and $e_{i,t}^{comm}$. Table A.5 reports the IV estimates. Again a negative relationship between the number of competitors and hosts' effort is present and it holds even after allowing for neighborhood-specific trends.

Table A.4: Evidence Supporting Assumption OC_2: Response Rate, $\bar{r}_{i,t}^{location}$, $e_{i,t}$

	Response rate	Response rate	Response rate
$\bar{r}_{i,t}^{location}\times 100$	0.0661		
	[0.0499]		
$e_{i,t}^{comm}\times 100$		0.213**	
		[0.107]	
$e_{i,t}^{check}\times 100$			0.307***
			[0.112]
Constant	0.967***	0.974***	0.974***
	[0.00478]	[0.000000875]	[0.00000110]
Listing FE	✓	✓	✓
Mean	0.972	0.972	0.972
R-squared	.0007727	.0001103	.0002375
N	68,371	68,371	68,371

Note: Only listings offering short-term lodging that enter the platform before September 2017 and exit after January 2018 are considered. Standard errors clustered by listing are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.5: IV Estimates of the Impact of Competition on Hosts' Effort

	$e_{i,t}^{check-in}$			$e_{i,t}^{comm}$		
	(1)	(2)	(3)	(4)	(5)	(6)
$\ln(L_{i,t}^{0.5})\ \times 10$	-0.841*			-0.748*		
	[0.434]			[0.409]		
$\ln(L_{i,t}^{1})\ \times 10$		-0.804**			-0.764**	
		[0.403]			[0.363]	
$\ln(L_{i,t}^{2})\ \times 10$			-0.663*			-0.531
			[0.384]			[0.349]
Constant	0.333*	0.418*	0.420	0.279	0.382*	0.316
	[0.195]	[0.232]	[0.269]	[0.183]	[0.208]	[0.244]
Listing FE	✓	✓	✓	✓	✓	✓
Snap FE	✓	✓	✓	✓	✓	✓
Mean	0.006	0.006	0.006	0.006	0.006	0.006
R^2	0.000144	0.000201	0.000400	0.000111	0.000159	0.000491
N	55,587	55,587	55,587	55,587	55,587	55,587

Note: Only listings offering short-term lodging that enter the platform before September 2017 and exit after January 2018 are considered. Standard errors clustered by listing are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.