Boston University School of Management

Boston University School of Management Research Paper Series

No. 2013-16

"The Rise of the Sharing Economy: Estimating the Impact of Airbnb on the Hotel Industry"

Electronic copy available at: http://ssrn.com/abstract=2366898

The Rise of the Sharing Economy: Estimating the Impact of Airbnb on the Hotel Industry[†]

Georgios Zervas
School of Management
Boston University

Davide Proserpio, John W. Byers
Computer Science Department
Boston University

February 12, 2014

Airbnb is an online community marketplace facilitating short-term rentals ranging from shared accommodations to entire homes that has now contributed more than ten million worldwide bookings to the so-called sharing economy. Our work addresses a central question facing the hospitality industry: to what extent are Airbnb stays serving as substitutes for hotel stays, and what is the impact on the bottom line of affected hotels? Our focus is the state of Texas, where we identify Airbnb's impact by exploiting significant spatiotemporal variation in the patterns of adoption across city-level markets. Using a dataset we collected spanning all Airbnb listings in Texas and a decade-long panel of quarterly tax revenue for all Texas hotels, we develop a nuanced estimate of Airbnb's material impact on hotel revenues. Our baseline estimate is that a 1% increase in Airbnb listings in Texas results in a 0.05% decrease in quarterly hotel revenues, an estimate compounded by Airbnb's rapid growth. To further isolate Airbnb's impact, we employ hotel segments that consumers are less likely to substitute for Airbnb stays as additional control groups. We find that the impacts are distributed unevenly across the industry, with lower-end hotels and hotels not catering to business travelers being the most affected. Finally, by simulating various regulatory interventions informed by current events, such as limiting Airbnb hosts to a single listing, we find only a moderate mitigating impact on hotel revenues.

[†]The authors thank the participants and organizers of SCECR'13 (`http://scecr.org/scecr2013/`), WISE'13 (`http://wiseconf.org`), and the seminar participants at Telefonica I+D Research, Barcelona and Technicolor Research, Paris for their helpful feedback on earlier drafts of this work. We thank Smith Travel Research (STR) for sharing data with us. We are also indebted to Flavio Esposito for motivating us to investigate Airbnb and for his contributions to our earlier research on the topic.

1 Introduction

Much ado has recently been made of the "sharing economy", in which broad segments of the population can collaboratively make use of under-utilized inventory via fee-based sharing. On the supply side, individuals can provide short-term rentals of vehicles they own that would otherwise be sitting idle, they can rent out spare rooms in their apartment or home, and now they can even rent out their pets.[1] On the demand side, consumers benefit from the sharing economy by renting goods at lower cost or with lower transactional overhead than buying or renting through a traditional provider. While sharing of goods and services is an age-old phenomenon, the key enabler to this marketplace operating at scale has been the rise of websites that facilitate far more of these transactions. In the first phase of this technology-driven boom, web sites like Craigslist enabled the *selling* of goods locally, primarily through the use of searchable listings, allowing suppliers to reach broader audiences at minimal cost, and for purchasers to browse a massive online inventory from their desktop. In the second phase, web sites such as Airbnb have addressed the much more challenging problem of the *sharing* of goods, where unlike selling a good, retaining (much of) the post-transaction value of the good is essential for the owner. Although we do not study this aspect of the sharing economy in this work, a key enabling technology has been the development of strong online reputation and signaling mechanisms. For example, online rental sites now prominently display the responsiveness of hosts to rental inquiries, which hosts can use to signal their quality relative to their competition. Also, many web sites promoting the sharing economy have embraced the use of online reviews – firms such as eBay pioneered bidirectional post-transaction reviews using ordinal star ratings; firms now encourage users to provide and publish detailed reviews of their experiences subsequent to each transaction.

Along with the rise of the sharing economy comes a host of open research questions. From the socioeconomic standpoint, tapping into and realizing the promise of the sharing economy comes with considerable attendant complexity, as new models challenge both existing business models and the underlying social fabric. Consider the case of Airbnb, a website enabling short-term rentals. In many markets, notably New York City, protectionist legislation promoted by the hotel industry forbids apartment owners from short-term rentals, and indeed, recent litigation has upheld these laws, effectively making Airbnb rental illegal.[2] Along a somewhat similar line, lack of regulation and oversight of gray-market transactions facilitated by the sharing economy may also lead to a shaking-out period during which regulatory

[1]See e.g., Borrow My Doggy - `https://www.borrowmydoggy.com/`

[2]See "Judge rules Airbnb illegal in New York City", May 2013, `http://money.cnn.com/2013/05/21/technology/innovation/airbnb-illegal-new-york/index.html`

action clamps down upon and limits the scope and sweep of the sharing economy.[3] Finally, under certain settings, the sharing economy imposes significant negative externalities upon non-participants. A canonical example is that of apartment dwellers being subjected to the whims of short-term renters, who may not ascribe to norms of noise, cleanliness, and public safety, and who provide no continuity in building and maintaining a sense of community.[4]

Ultimately, the sharing economy can and should be viewed through the lens of social welfare, which we take to be the standard interpretation: the sum of individual utilities over a society. Clearly, there exist beneficial transactions enabled through the sharing economy that have no negative externalities, benefit both transacting parties, and thus provide positive utility. Such a transaction manifests itself as a net gain in social welfare. Proponents of the sharing economy opine that a vast number of hypothetical transactions are of this form, and thus, the ability to unlock the sharing economy would provide a great boon to social welfare (see, *e.g.,* Botsman (2012)). Detractors, as well as entrenched interests, argue that the sharing economy presents a societal risk, and while under certain scenarios, a properly regulated sharing economy may add value, the net gain to social welfare in the long run could be insubstantial.[5]

In this paper, we empirically examine the impact that the rise of the sharing economy, and specifically the emerging market for short-term rentals, has on the hospitality industry. We investigate whether and to what extent, stays in the short-term apartment and home rental market displace hotel stays. Using data we collected from Airbnb.com on over 22,000 stays in the state of Texas over the five year period from 2008 to 2013, and quarterly hotel revenue tax data from over 4,000 hotels in Texas dating back to 2003, we show that Airbnb penetration is negatively correlated with hotel revenue, and that lower-end hotels incur most of the financial impact. In particular, due to the significant variability in both the temporal rate and the spatial density of Airbnb penetration, as well as the geographic specificity of both our hotel and Airbnb datasets, we are able to treat Airbnb market entry as a variable intervention in space and time against the hotel tax revenue dataset. Our analysis also accounts for differences in consumer behavior across regions (municipalities) of Texas through a rich set of control variables, and exploits the panel nature of our dataset by incorporating hotel fixed effects. Our preferred specification allows us to project that, in Texas, an additional 1% increase in the size of the Airbnb market will result in a 0.05%

[3]See "Sharing Verboten: Berlin Puts Kibosh on Airbnb and Co.", August 2013, `http://www.spiegel.de/international/business/berlin-to-penalize-short-term-rental-companies-like-airbnb-in-fall-a-916416.html`.

[4]See "Is Airbnb Bothering the Neighbors?", September 2012, `http://betabeat.com/2012/01/is-airbnb-bothering-the-neighbors/`.

[5]See "The Rise Of The Renting And Sharing Economy Could Have Catastrophic Ripple Effects", August 2013, `http://www.businessinsider.com/rise-of-the-renting-and-sharing-economy-2013-8?op=1`

decrease in total hotel revenue.

To provide additional robustness to this finding, we then move to study our hypothesis that Airbnb does not affect all hotels in a given municipality equally. Using a design similar to a difference-in-difference-in-differences analysis, we incorporate high-end hotels, and hotels that cater to business travelers as additional control groups. We find that Airbnb disproportionately impacts lower-end properties, as one might expect, given the nature of rentals on Airbnb, typically with fewer amenities and services that higher-end hotels provide. Through a similar analysis, we find that the impact of Airbnb also falls disproportionately on hotels with little or no conference space, our proxy for the extent to which the hotel caters to business travel, as we hypothesized.

Finally, we perform a series of counterfactual simulations to estimate Airbnb's impact on hotel revenue. We focus attention on the expected effects of possible regulatory interventions informed by laws and regulations related to illegal hotels, as well as Airbnb's terms and conditions. We first estimate the effects of disallowing the rental of non-shared accommodations, then consider limiting Airbnb hosts to one listing each. We find that the improvement of hotel revenues is more significant when the former regulation is applied as opposed to the latter, but neither regulation individually eliminates a significant proportion of Airbnb's impact on hotel revenues.

2 Related Work

Our work contributes to the small but expanding literature on substitution between online and offline markets, as firms like Airbnb can be viewed as providing enabling technology that facilitates suppliers of niche inventory to bring their products to market, online. In contrast to offline markets, Airbnb provides sufficient reach and sufficiently low cost of revenue for individuals to profitably list their remnant inventory. As such, our study of Airbnb can be viewed as investigating the consequences of lowering the barrier to entry for online suppliers in relation to traditional suppliers.

Brynjolfsson et al. (2009) examine the role of product popularity with respect to consumers' choices among bricks-and-mortar retailers, shopping catalogs, and electronic commerce stores. They find that substitution between online and offline channels is less intense for niche products which are generally less likely to be available offline. Forman et al. (2009) consider the role of consumers' physical locations, and find that consumers who live closer to bricks-and-mortar stores have a stronger preference for them; furthermore, these consumers are less sensitive to online discounts. Similar to aspects of these works, our study and identification strategy focuses on settings where 1) the online offering arises subsequent to an established offline market presence, 2) there is significant observable variability with respect to the emergence of the online offering, and 3) geographic variability in the offerings plays a key role with respect to consumer choice. However, our work differs substantially from these studies, as they consider the case of identical goods, whereas we consider the case where the online products are different, often distinctively so, from traditional offerings, and cannot be considered pure substitutes. Moreover, we consider the case where a new technology platform enables a wide range of micro-suppliers to enter the market, as opposed to prior work's focus on the emergence of a monolithic online alternative to retail stores.

Prior work has also examined substitution between online and offline advertising and retailing channels. The study by Goldfarb and Tucker (2011a) exploits cross-state regulatory differences to show that in states where lawyers are prohibited from contacting potential clients by traditional mail, costs-per-click for online search engine ads are substantially higher. A second study by Goldfarb and Tucker (2011b) exploits local variation in offline alcohol advertising bans, and finds that while alcohol advertising bans reduce alcohol purchase intent, this reduction is significantly smaller for consumers who are exposed to online advertising. Finally, Goldfarb and Tucker (2011c) summarize the progress on measuring substitution between online and offline advertising, and discuss the methodological challenges facing researchers who wish to identify such patterns, many of which also arise in our work.

A different body of relevant related work considers online technologies which enable suppliers of niche goods and services, or consumers with remnant inventory, to reach broader audiences. For example, a large number of recent studies have focused on the impact of Craigslist – a website featuring free online classified ads – on the newspaper industry. Seamans and Zhu (2013) estimate the effect of Craigslist's market entry on several newspaper performance metrics. They find that in the face of increasing competition by Craigslist, newspapers with greater reliance on classified ad revenue responded by reducing their ad rates, and by increasing their subscription prices more than newspapers whose revenues were less reliant on advertising. On the buy side, they find that this increase in competition translated to approximately \$5 billion worth of savings that accrue to advertisers. Along similar lines, Kroft and Pope (2013) estimate that Craigslist's entry resulted in a 7% reduction in the volume of classified ads appearing in newspapers during the period between January 2005 and April 2007. Further, they estimate that Craiglist's entry caused a decrease in the rental vacancy rate by approximately 1%. In a different line, Chan and Ghose (2011) provide evidence that the entry of Craigslist has led to a significant increase in the rate of new sexually transmitted disease cases by facilitating the posting of casual encounter ads. Our work shares a methodological trait with these studies: all of them rely on the temporal and geographic variation in Craigslist's entry to identify its effect. We exploit similar variation in the patterns of Airbnb's diffusion to measure its impact on hotel revenues.

Our work also contributes to the literature studying the impact the external shocks on the tourism and the hospitality industry. Much of the prior work though, has centered on demand shocks. For example, O'Connor et al. (2008) study the impact of terrorism on tourism in Ireland, and Baker and Coulter (2007) estimate the impact of the 2002 and 2005 terrorist attacks in Bali on the islands' vendors. Most closely related to our work is the study of Kosova and Enz (2012) who examine the adverse effects of the 9/11 attack and the 2008 financial crisis on hotel performance. Their findings on the 2008 financial crisis are directly relevant to our work since the events leading to the crisis overlap with Airbnb's initial entry into the Texas market, a factor we must correct for. Overall, they find that hotel revenues recovered gradually, returning to pre-crisis levels of performance around the beginning of 2009. Among different hotel segments, upmarket properties were those that suffered most, but also the ones that rebounded fastest. In contrast to these works, our research considers impacts from the supply side, as Airbnb rentals increase the supply of accommodation alternatives available to travelers. Moreover, unlike one-time, industry-wide demand shocks, the impact introduced by Airbnb is gradual and exhibits geographic variability, adding considerable nuance to appropriate identification strategies.

Finally, our work is related specifically to research that investigates the economic impacts from the sharing economy. A number of studies have been conducted on the adoption and effects of car-sharing. Cervero et al. (2007) study car-sharing in San Francisco using survey analysis methods to find that car-sharing is associated with significant decreases in miles traveled, gasoline consumption, and car ownership. In a similar survey of approximately 6,000 North American households participating in a car-sharing program, Martin et al. (2010) find that the average number of vehicles owned per household dropped by nearly 50%. Our analysis uncovers similar substitution patterns in the context of travel accommodation. We also find a large number of opinion pieces on accommodation sharing, for example in the popular press and on blogs, but little in the way of academic literature. Our closest comparison point are a set of short studies, commissioned by Airbnb, which argue that the Airbnb business model is complementary to the hotel industry and quantify the substantial net economic benefit to cities that Airbnb renters provide.[6] While our work is related to these

[6] See:

1. "Study Finds that Airbnb Hosts and Guests Have Major Positive Effect on City Economies", November 2012, `https://www.airbnb.com/press/news/study-finds-that-airbnb-hosts-and-guests-have-major-positive-effect-on-city-economies`.
2. "New Study: Airbnb Community Contributes €185 million to Parisian Economy", June 2013. `https://www.airbnb.com/press/news/new-study-airbnb-community-contributes-185-million-to-parisian-economy`.
3. "New Study: Airbnb Community Contributes $130 Million to Berlin Economy", Septem-

studies, we apply a more sophisticated methodology and segmentation analysis, resulting in conclusions that are both different and more nuanced. For example, our empirical strategy results in our quantitative identification of the non-negligible revenue impact that Airbnb has had on the Texas lodging industry, broken out by market segment.

3 The rise of Airbnb

Airbnb defines itself as *"a social website that connects people who have space to spare with those who are looking for a place to stay"*, and exemplifies a community marketplace. Airbnb hosts list their spare rooms or apartments, establish their own nightly, weekly or monthly price, and offer accommodation to Airbnb guests. Airbnb derives revenue from both guests and hosts for this service. They charge guests a $9-12\%$ service fee every time a reservation is booked, depending on the length of the reservation, and they charge hosts a 3% service fee to cover the cost of processing payments.

Airbnb's business model currently operates with minimal regulatory controls in most locations, and as a result, hosts and guests both have incentives to use signalling mechanisms to build trust and maximize the likelihood of a successful booking. To reinforce this behavior, Airbnb has built an online reputation system that enables and encourages each guest and host to leave a review upon completion of a stay. Guests use star ratings to rate features of their stay, *e.g.*, cleanliness, location, and communication, while both guests and hosts may provide other information about aspects of the stay, including personal comments.

Since its launch in 2008, the Airbnb online marketplace has experienced very rapid growth, with more than four million guests and over ten million nights of cumulative bookings worldwide at the end of 2012. The site is now being used by over 50,000 renters per night and had a market cap of $2.5 billion after its most recent funding round, late in 2012.[7]

4 Datasets and Data Preprocessing

For our study, we collect and combine data from various sources including the Airbnb website, the Texas Comptroller Office, Smith Travel Research (STR), county demographics from the U.S. Census Bureau, and the Current Population Survey (CPS) from the U.S. Bureau of Labor Statistics (BLS).

ber 2013. `http://publicpolicy.airbnb.com/wp-content/uploads/2013/09/Berlin-Airbnb-economic-impact-study.pdf`.

[7] http://go.bloomberg.com/tech-deals/2013-03-10-the-missed-airbnb-investment-now-worth-250-million/

4.1 Airbnb Data

To estimate the extent of Airbnb's market entry, we collected consumer-facing information from Airbnb.com on the complete set of users who had listed their properties in the state of Texas for rental on Airbnb. We refer to these users as *hosts*, and their properties as their *listings*. Each host is associated with a set of attributes including a photo, a personal statement, their listings, guest reviews of their properties, and Airbnb-certified contact information. Hosts who supply Airbnb with a copy of a government-issued photo ID receive a prominently displayed "Verified ID" badge. Similarly, each listing displays attributes including location, starting price, a brief textual description, photos, capacity, availability, check-in and check-out times, cleaning fees, and security deposits. Figure 1 displays a typical Airbnb listing, and Figure 2 displays a typical Airbnb user profile. Our collected dataset contains detailed information on $5,994$ distinct hosts and $7,361$ distinct listings.

Airbnb provides a number of tools for guests and hosts to build their online reputations. Most prominent among these tools is an archival reputation system that solicits, records, and displays crowdsourced reviews from Airbnb users. After each stay, both the guest and the host are asked by Airbnb to provide a review about their experience, which are then published on the publicly available host and guest pages, respectively. We collected all user reviews from transactions involving the Texas listings, but for this paper, make use only of the subset of 22,650 reviews written by guests for the Texas listings for transactions between 2008 and Q2 2013.[8]

From our collected data, we are able to quantify certain measures of spatiotemporal penetration of Airbnb. Ideally we would use bookings data to estimate penetration, but individual transaction values on Airbnb are not available to us. Instead, for any given region in space, *e.g.,* city, ZIP code, state; and for any given point in time, we define penetration to be the number of distinct Airbnb listings that have cumulatively entered the market in that region up to that point in time. We approximate the entry date of individual listings by using the (prominently displayed) date their owners became Airbnb members. A clear limitation of this measure is the lack of direct correspondence between the supply of Airbnb listings and consumer demand for these listings. This challenge is not unique to our setting. For example, to analyze the impact of Craigslist's penetration, Seamans and Zhu (2013) and Chan and Ghose (2011) rely on similar proxy measures such as the number of classified ads available on the website, or, more simply, the appearance of the first classified ad in a

[8]The Airbnb review corpus has some distinctive differences from other collections of online reviews. Unlike forums like eBay, Airbnb does not use star-ratings or other summaries to rank users; it simply publishes the review content. Also, users may participate on both sides of the market, and thus may build both a reputation as a host and as a guest.

region. In Section 7, we perform a series of robustness checks with different proxy measures to investigate the sensitivity of our results to this choice.

Separately, we must choose an appropriate level of geographic aggregation. Here, our data is suitably granular (with location accuracy to roughly 100 meters) to permit analysis at many different scales. Our use of city-level granularity in our primary specification is driven by the observation that a city is the largest geographic unit within which we reasonably expect to see significant substitution patterns between hotels and Airbnb properties. We anticipate and assume that the fraction of travelers who will substitute a hotel for an Airbnb room in a different city is insignificant.

4.2 Hotel Revenue Data

The dependent variable we use in our analysis of the impact of Airbnb is hotel revenues. We obtained quarterly tax panel dataset for hotels in Texas from the website of the Texas Comptroller Office, which publishes tax reports for individual hotels.[9] In addition to basic information including hotel name, address, and capacity, the panel data includes all hotel room receipts, broken out as taxable and non-taxable receipts, by quarter.[10] The raw dataset spans the period between the first quarter of 2003 and the second quarter of 2013 and contains approximately 330,000 records. Interestingly, according to Texas law,

> "... a hotel is considered to be any building in which members of the public rent sleeping accommodations for $15 or more per day."[11]

For this reason, Airbnb properties complying with the Texas tax code appear in this tax dataset. This is evident from Figure 3, which plots the quarterly number of unique tax-paying properties in Austin broken down by capacity, *i.e.*, maximum occupancy. We conjecture that the rapid increase in low capacity properties starting in 2008 is related to Airbnb's entry into the Texas market at the same time. To exclude non-hotel properties from our analysis, we cross-reference the tax dataset with the U.S. hotel census data provided to us by STR. The STR census includes all U.S. hotels and contains a rich attribute set for each hotel including its opening date, price segment, capacity, operation (chain vs. independent), and geographic location. In total, the STR dataset contains information on 4,673 hotels in the state of Texas. After linking the STR census dataset with the Texas tax dataset, we obtain high-confidence matches for a panel of 4,006 properties (86% of the Texas hotels in the STR census), and exclude the rest. Table 1 provides summary statistics of the variables we use

[9] Available at `http://aixtcp.cpa.state.tx.us/hotel/hotel_qtr_all_srch.php`

[10] Differences due to non-taxable receipts are generally small. We use total room receipts as our measure.

[11] `http://www.window.state.tx.us/taxinfo/hotel/faqhotel.html`

in our analysis from 2003 through June 2013. Each pair of entries are computed by taking the average and standard deviation of each variable for all observations in our dataset falling within a given year.

Finally, we assemble a set of control variables derived from publicly available data sources. We obtain quarterly unemployment data at the city level, annual demographic information at the county level, and quarterly employment in the accommodation industry at the county level from the BLS at `bls.gov` and the U.S. Census Bureau at `census.gov`. We also adjust for inflation using quarterly Consumer Price Indexes (CPIs) reported by the BLS.

5 Empirical Strategy

Airbnb has seen widely varying degrees of traction within different local, regional and international markets, both with respect to initial market entry and the rate at which it has been adopted within markets. For example, consider Figure 4, which depicts the current extent of market penetration both of Airbnb properties and hotels within the state of Texas (top panels), and within the county encompassing the state capital, Austin (bottom panels). Unlike hotels, which have coverage throughout the state and pockets of local density, such as in downtown Austin, Airbnb has spotty coverage at best throughout the state, but broader coverage across metro areas including suburbs and exurbs. Table 2 reveals that patterns of Airbnb adoption over a five year period in the ten most populous cities in Texas are themselves diverse, with several cities experiencing early adoption and rapid growth and others experiencing minimal impacts. Our empirical strategy exploits this variability to identify the impact of Airbnb's rise on hotel revenues.

Our base specification takes the following form:

$$\log \text{Hotel Revenue}_{it} = \beta \log \text{ Airbnb Listings}_{it} + x'_{it}\gamma + \alpha_i + \tau_t + \epsilon_{it}. \qquad (1)$$

We construct the dependent variable of our specification by applying two transformations to quarterly hotel revenues, indexed both by quarter t and hotel i. First, since our data spans a decade of observations, we adjust hotel revenues for inflation, correcting all revenue to January 2003 dollars using the All Prices Consumer Price Indexes (CPIs). Second, we opt for a logarithmic transformation based on the observation that the distribution of hotel revenues is highly skewed, as shown in the left panel of Figure 5. By contrast, as shown in the right panel of the same figure, the logarithmic transformation better approximates a normal distribution. An additional advantage of the logarithmic transformation is that it constrains the predictions of our model to positive revenue values.

The coefficient of interest is β, which we interpret as the percentage change in hotel revenues caused by a 1% increase in the number of Airbnb listings. The vector x_{it} contains a set of time-varying control variables (x' denotes the transpose of this vector). In x_{it} we include: the total number of hotel rooms measured at the city level (we expect that the supply of hotel rooms will also affect revenues), as well as unemployment rate, population, and the number of employees in the leisure and accommodation industry measured at the county level. We apply a logarithmic transformation to all controls except for unemployment, which is already expressed as a rate. To further control for pre-existing trends across cities we also include city-specific quadratic time-trends in our set of controls. A concern with the inclusion of city-specific time-trends is that they can be confounded with hotels' response to Airbnb. Wolfers (2006) examines this issue in detail, and highlights problems that arise when time-trends are estimated on a small number of pre-intervention observations. Fortunately, our dataset covers a long pre-Airbnb period from 2003 to 2008, alleviating these concerns. Our specification also includes hotel fixed effects (variables α_i) to control for unobservable time-invariant hotel properties, and year-quarter fixed effects (variables τ_t) to control for concurrent effects common across all hotels (*e.g.*, the 2008 financial crisis.) Finally, we cluster standard errors at the hotel level to account for autocorrelation in our data (see *e.g.*, the treatment recommended by Bertrand et al. (2004)).

We report our results in Table 3. We begin with the simplest specification lacking any controls x_{it}, the results of which are reported in the first column. We find that a 1% increase in Airbnb listings is associated with a statistically significant 0.04% ($p < 0.01$) decrease in quarterly hotel revenues. We then proceed to check the robustness of this relationship by successively incorporating the set of controls x_{it} elaborated above. Our estimations with these added controls, reported in columns two through five of Table 3, remain similar in directionality; moreover, the resulting coefficients of our control variables have reasonable direction and magnitudes. For example, the coefficients of unemployment rate and the number of hotel rooms are both negative and significant, while the number of employees in the accommodation sector is positively correlated with hotel revenues, lending additional credibility to our model specification. Ultimately, in our preferred specification (Column (5)), a 1% increase in Airbnb listings results in the association with largest magnitude, a 0.05% ($p < 0.01$) decrease in quarterly hotel revenues.[12]

Placing our results in context, we observe that our preferred specification in Table 3 suggests that a percentage increase in the supply of hotel rooms in Texas is associated with

[12]The number of observations is smaller in this specification because our dataset lacks the number of employees in the accommodation industry for a small number of counties. An alternative specification where we impute the missing values by substituting with the Texas-wide average produced almost identical results. Furthermore, these missing counties account for a very small fraction of the Airbnb listings in our dataset.

a roughly 0.29% decrease in Texas hotel revenues. Therefore, even at the current modest rate of Airbnb penetration, a percentage point increase in the supply of Airbnb listings has roughly a sixth of the impact of a comparable percentage increase in hotel rooms.[13] This makes intuitive sense, as any Airbnb stay that substitutes a hotel room stay equates to lost revenue for the hotel industry, whereas increased hotel room supply merely provides diminished revenue through increased price competitiveness. But this currently substantial impact is all the more striking in light of the fact that Airbnb continues to grow rapidly. For example, the number of Airbnb listings in Austin more than doubled between 2011 and 2012 (see Table 2). Furthermore, larger markets in Texas such as Houston and Dallas appear to have ample room for Airbnb growth, due to their large population, and relatively low Airbnb penetration. Therefore, our results suggest that market entrants such as Airbnb can pose a legitimate competitive threat to the hospitality industry going forward.

5.1 Further isolating the Airbnb effect

While Airbnb can provide an alternative to hotels, one can hardly expect it to be a perfect substitute for all travel needs. Here, we attempt to further isolate the impact of Airbnb on hotel revenues by focusing on two hotel segments that we argue should be less vulnerable to Airbnb's entry: high-end hotels, and hotels catering to business travelers. Motivated by this intuition, we estimate a set of regressions conceptually analogous to a difference-in-difference-in-differences (DDD) design. Observe that our base specification in Equation 1 can be interpreted as a difference-in-differences estimate where each unit i acts as its own control. Our goal is to arrive at a more robust estimate of the impact of Airbnb by employing those hotels least affected by Airbnb's entry as additional controls (a third difference).

5.1.1 High-end hotels as a control group

On the basis of price, Airbnb inventory is most comparable to the Budget, Economy, and Midprice segments of the hotel inventory in Texas. Our comparison is furnished in Figure 6. The left panel of this figure plots 2012 average daily rates for all hotels in Texas, and for the top-5 cities broken down by price segment, plotting reservation data that we obtained from STR and `hotels.com`. The right panel of Figure 6 presents the average offered nightly rates for Airbnb properties in Texas in 2013.

Based both on this price comparison and the fact that upmarket hotels provide amenities (*e.g.,* pools, conference rooms, concierge) to travelers that typical Airbnb rentals do not, we

[13]In absolute terms, the total supply of Airbnb listings in Texas is currently about 7,300, while the number of hotel rooms is approximately 360,000 (as per STR hotel census.)

hypothesize that the impact of Airbnb will differ substantially by hotel price segment. We expect the impact to be disproportionately larger for lower price tiers. To investigate this hypothesis, we first estimate separate regressions for hotels by price segment. The results, shown in the first five columns of Table 4, largely support our hypothesis. While the impact is negative across segments, the magnitude is largest for lower price tiers, while for hotels in the Luxury classification, the estimated Airbnb effect is not significant.

Next, we estimate the following model using our complete dataset:

$$\begin{aligned}\log \text{Hotel Revenue}_{it} = &\beta_1 \text{log Airbnb Listings}_{it} \\ &+ \beta_2 \text{log Airbnb Listings}_{it} \times \text{Price Segment}_i \\ &+ x'_{it}\gamma + \alpha_i + \tau_t + \epsilon_{it}. \end{aligned} \tag{2}$$

We report our results in the rightmost column of Table 4. Both the main effect (coefficient β_1), and its interactions with the Budget, Economy, and Midprice hotel price segments (coefficients β_2) are statistically significant. Furthermore, the interaction coefficients, which are the ones of interest in this analysis, have the expected direction and magnitude: using Luxury hotels as the reference group, we find that the negative impact of Airbnb is amplified as we move to lower price tiers, with Economy hotels being most impacted. Upscale hotels on the other hand, are not significantly affected compared to the Luxury segment. According to this analysis the impact of Airbnb remains significant and substantial on lower-end hotels even once we control for revenue trends common across different price tiers.[14]

From a managerial standpoint, this result has direct import: even though lower-end hotels in Texas account for a disproportionately small amount of revenue as compared with upmarket hotels, they nevertheless bear the brunt of the impact of the market entry of Airbnb. Our evidence suggests that consumers are increasingly substituting Airbnb stays for lower-end hotels in Texas, possibly identifying the former as offering better value at a similar price point. While this increased competition affords consumers greater choice, it also places lower-end hotels in regions with high Airbnb penetration at greater risk.

5.1.2 Business travel hotels as a control group

A second distinction that relates to the substitutability between hotel and Airbnb stays regards the extent to which hotels are used for business travel. With most business travel traditionally arranged through employers, an Airbnb stay may be more difficult to substitute.

[14] A concern with this analysis is our choice of Luxury hotels as the control group. As is evident in Figure 8 their average quarterly revenues are significantly higher compared to lower price segments, with more pronounced inter-quarter variation. As a robustness check, we repeated the same analysis using Upscale hotels as the control group, omitting Luxury hotels altogether. Our results remained essentially unchanged.

Also, business travelers may make greater use of those hotel amenities not typically provided in an Airbnb stay. Concretely, consider hotels with conference and meeting facilities. This amenity primarily serves business travelers, drives accommodation revenue through stays by conference attendees, and is a service rarely provided with an Airbnb stay. For these reasons, we anticipate that hotels with larger conference and meeting spaces will see less of an impact on their revenues as a direct consequence of Airbnb's entry. This observation motivates a second DDD design, this time employing conference hotels as a control group. Our hypothesis is that the most affected segment of the hotel industry are those hotels in areas with high Airbnb penetration but which are less likely to attract business travelers.

Our STR hotel dataset spans Texas hotels with wide variety in meeting space size, facilitating this comparative analysis. Approximately 10% of the hotels in our dataset have at least 10,000 square feet of meeting space. For reference, spaces of this size can usually accommodate hundreds of people. We begin our analysis by separately estimating the impact of Airbnb on hotels with successively higher cutoffs for meeting space: no meeting space, at least 500 square feet of meeting space, at least 1,000 square feet, and at least 2,000 square feet. We present our results in the first four columns of Table 5, which shows that as we limit our analysis to hotels with larger cutoffs, the impact of Airbnb becomes progressively smaller. We then estimate the following model on our complete dataset:

$$
\begin{aligned}
\log \text{Hotel Revenue}_{it} =& \beta_1 \text{log Airbnb Listings}_{it} \\
&+ \beta_2 \text{log Airbnb Listings}_{it} \times \text{log Meeting Space}_i \\
&+ x'_{it}\gamma + \alpha_i + \tau_t + \epsilon_{it}. \qquad (3)
\end{aligned}
$$

The coefficient of interest in this model is β_2, for the term representing the interaction of *log Airbnb Listings* and *log Meeting Space*. According to our hypothesis, hotels with conference facilities should experience less of an impact due to Airbnb and hence β_2 should be positive. Our results, displayed in the final column of Table 5, are consistent with this hypothesis: a unit increase in log Meeting Space reduces the magnitude of the Airbnb effect by 0.002 ($p < 0.01$). Meanwhile, the coefficient of the main effect, β_1, remains negative and statistically significant.

In summary, we argue that whereas differences in hotel revenues between markets with varying degrees of Airbnb adoption could potentially arise due to unobserved factors unrelated to Airbnb, the DDD methodology applied here is much less prone to that line of argumentation. By contrasting hotel revenues along the dimension of meeting space in addition to existing controls, we have further isolated the effect of Airbnb on a hotel segment that we hypothesized should be most vulnerable to the rise of Airbnb: hotels that are located in areas with high Airbnb adoption rates *and* are less likely to attract business travelers.

6 Simulation of Counterfactual Outcomes

In the previous section we estimated Airbnb's overall impact on hotel revenue, and the variation of impact across different hotel price categories. However, the point estimates we derived by category (such as changes in Economy hotel revenue with Luxury hotels as a control group) provide limited insight on counterfactual outcomes. In this section, we employ statistical simulation to convey our results more intuitively while incorporating the uncertainty built into our estimates of model parameters (see King et al. (2000)). We first focus on Austin – where Airbnb has been most popular – and ask: what would the revenues of Austin hotels have been if there had been fewer Airbnb listings available? We also consider counterfactual outcomes in which Airbnb grew even faster than observed, to derive insights into hypothetical future scenarios. We finally consider the effect of hypothetical regulatory interventions motivated by current events.

As a first step towards answering this question we recast the fixed-effects model in Equation (2) as a random-effects model where we treat the α_i as random normal terms with unknown variance (to be estimated). This random-effects specification better serves our simulation goal since, unlike its fixed-effects counterpart, it can be used to make predictions for hotels outside our sample, including the average hotel. The main shortcoming in random-effects models is that the α_i term is assumed to be orthogonal to observed variables. We address this concern in two ways. First, we incorporate hotel capacity, price category, and city as additional controls in our model. Second, following standard practice (see Mundlak (1978)), we explicitly allow for correlation between observables and unobservables by decomposing hotel random effects to $\alpha_i = \delta_i \bar{x}_i + \mu_i$, where μ_i is a zero-mean random normal term, and $\bar{x}_i$ are the group-means of all time-varying variables in our model.[15] The estimates of the random effects model, omitted for brevity, remain virtually unchanged compared to the corresponding fixed-effect estimates shown in the rightmost column of Table 4.

Next, we use the random-effects model to compute an expectation for the quarterly revenues of the average hotel for different price categories under different scenarios of Airbnb penetration. The algorithms we use are described in detail by King et al. (2000). We summarize them here for completeness. First, recall that our model can be described in terms of the following parameters: the coefficients β, the variance σ_α^2 of hotel random effects α_i, and the variance σ_ϵ^2 of the idiosyncratic error term ϵ_{it}. Additionally, let x be a vector of

[15] This correction is easily implemented by incorporating the group means $\bar{x}_i$ as regressors in our model.

explanatory variables appropriately chosen to describe the setting for which we are computing an expected value (x, for example, contains as one element the number of Airbnb listings, and as another the price category of the hotel.) To obtain an expectation for the average hotel we set α_i to zero. Then, to compute a single expected value we do the following:

1. Draw a parameter vector $(\hat{\beta}, \hat{\sigma_\epsilon})$ from the sampling distribution of our model's estimators.[16]

2. Using the vector $(\hat{\beta}, \hat{\sigma_\epsilon})$ we produce 1000 samples from a normal distribution $N(\hat{\beta}x', \hat{\sigma}_\epsilon)$.

3. Since the dependent variable is expressed in a log-scale we apply an exponential transformation to the above samples to convert them to dollar values, and take their sample mean to arrive at a single expected value.

To approximate the full distribution of the expected value of a hotel's quarterly revenues we repeat the above process 1000 times, and use these samples to compute quantities of interest such as the mean expected revenue, and its standard error.

A final technical point merits discussion. When we perform simulations of alternative levels of Airbnb penetration we only manipulate the explanatory variables associated with the interaction coefficients in Equation 2, while we keep the explanatory variable associated with the main effect at its observed levels. This leads to more conservative estimates of Airbnb's effect, and captures the intuition that we measure changes in revenue above and beyond any changes in the revenue trends of our Luxury hotel control group.

For our simulations, we consider two sets of counterfactuals. The first is motivated by an alternative reality in which Airbnb did not come to exist. An obvious way to proceed with simulating this scenario would be to assume that the number of Airbnb listings would be zero. However, we believe that such an assumption is unrealistic. In a world without Airbnb, some consumers would have likely used one of many other websites that serve a similar purpose, such as Vacation Rentals by Owner (VRBO), HomeAway, or even Craigslist, to make their properties available for rental. At the same time, given the inability of incumbent sites to match Airbnb's growth, we conjecture that the fraction of listings that would have found their way online if Airbnb did not exist is small. While acknowledging that this fraction is difficult to estimate, we propose simulating two counterfactuals where 1% and 10% of Airbnb's inventory is made available online via alternative sites. Our first set of simulations is shown in Figure 9. For each hotel category other than the control group (Luxury), for which we assume the Airbnb effect to be zero, we plot the mean expected value of the first

[16]The coefficients of the model follow a multivariate-normal sampling distribution, while the error term follows an inverse scaled chi-square distribution.

difference in hotel revenues by quarter between observed levels of Airbnb penetration and the counterfactual of 1% Airbnb penetration. The solid lines can be interpreted as the expected financial impact of Airbnb on the average hotel in Austin over time (or, alternatively the expected gain for the average hotel if a large fraction of Airbnb inventory is eliminated). The grey bands are 95% confidence intervals. These plots depict a negative trend in revenue coincident with the arrival of Airbnb in 2008 for each of the three lowest hotel categories, while that of Upscale is inconclusive. For the three lowest categories, the revenue gap appears to widen over time, most prominently for the Economy segment, due to the sharp rise of Airbnb listings. The difference in quarterly revenues for the average Economy hotel in Austin grows to an absolute value that exceeds $50,000 in expectation in early 2013. The plots also highlight a limitation of our modeling strategy: due to the logarithmic nature of our model, small absolute changes in Airbnb inventory during Airbnb's early stages have a seemingly disproportionately high effect on hotel revenues. However this modeling choice is not arbitrary; we chose it intentionally to incorporate diminishing hotel revenue impacts as the marginal supply of Airbnb rooms increases.

In Figure 10, we apply the same methodology to simulate revenue differences for all of 2012, the most recent year for which we have complete data. Yearly estimates are of interest because unlike quarterly estimates, they are not sensitive to seasonal variation (see, *e.g.* Figure 8). The two darkest bars correspond to differences between observed levels of Airbnb penetration and the 1% and 10% alternatives. We estimate that the difference at the 1% penetration level is $-\$84,000 \pm \$28,000$ for an average Budget hotel; $-\$250,000 \pm \$61,000$ for Economy; and, $-\$200,000 \pm \$142,000$ for Midprice. For an average Upscale hotel, the difference is not significantly different from zero. If we assume 10% Airbnb penetration, these estimates become substantially lower: $-\$40,000 \pm \$14,000$ for Budget, $-\$119,000 \pm \$32,000$ for Economy, and $-\$98,000 \pm \$75,000$ for Midprice. These estimates are reflective of significant substitution patterns between Economy hotels and Airbnb accommodations.

Next, we turn our attention to predicting the impact of Airbnb's growth in the future. While it is beyond the scope of this work to predict the rate at which Airbnb will grow, our methods can readily be extended to analyze the counterfactual setting in which Airbnb adoption doubles, while keeping other variables in our analysis constant at their latest observed values. Here, we return our focus to the state-wide Airbnb impact, by forming a prediction for every hotel in Texas. To proceed, we first obtain predictions for the random components α_i for all hotels in our dataset. We then use these predicted α_i's in our simulation as above, enabling us to compute first-differences in quarterly revenue between current rates of Airbnb penetration and doubled rates of penetration. We assume that under this Airbnb-doubling scenario there will be no new Airbnb listings in areas that currently have no

Airbnb listings. In percentage terms, a doubling of Airbnb amounts to an additional revenue shortfall of −2.1% (with a standard deviation of 0.36%) for Budget hotels, −2.6% (0.32%) for Economy hotels, and −0.9% (0.32%) for Midprice hotels. The effect is insignificant for Upscale properties.

We next turn our attention to scenarios motivated by the regulatory and legal headwinds that Airbnb currently faces. For example, in NYC, Airbnb is currently being challenged by litigators who allege that Airbnb accommodations not shared by a landlord are in violation of city law, and in many cases, the leases that tenants signed. In related legal action, regulators are targeting owners who have multiple units on offer, arguing that those owners are running illegal and unregulated hotels. These considerations are encapsulated in our second set of simulations, in which regulators successfully either eliminate non-shared rooms, or limit Airbnb landlords to a single unit, respectively. To put these various effects in perspective, in Austin there are (as of Q2 2013) 5,468 listings, of which 3,979 are managed by single-property hosts, while the remaining 1,489 listings are managed by 594 hosts. Segmenting Airbnb listings by type, the Austin market consists of 3,771 entire apartments and 1,697 private or shared rooms.[17] The results of these simulations are also displayed in Figure 10 (using the two lightest shades), and they can be interpreted as the expected hotel revenue loss in 2012 incurred due to entire Airbnb apartments, and due to Airbnb properties belonging to hosts who manage multiple listings. Alternatively, the absolute values of these numbers can be interpreted as the expected gains in revenue following successful implementation of the two regulatory interventions we consider. Of the two interventions, the most significant expected gains for hotels are realized when eliminating whole apartments from Airbnb inventory. We estimate that the expected yearly revenue gain for the average Economy hotel from eliminating entire apartments from Airbnb inventory is $\$56,000 \pm \$22,000$. The expected gain for the average Economy hotel when limiting Airbnb hosts to a single property each is smaller: $\$11,000 \pm \$18,000$.

When comparing the expected effects of these regulatory interventions, we should take into account two facts our simulations do not capture. First, the number of Airbnb entrepreneurs flouting the law by renting out multiple properties is much smaller than the number of hosts renting out entire apartments, making regulation directed at them easier to enforce.[18] Second, according to our data, while multi-unit hosts manage 27% of Airbnb

[17] In NYC, which is the focal point of conversation regarding regulation, there are 18,421 listings, of which 11,309 are managed by single-property hosts, while the remaining 7,112 are managed by 1,869 hosts. In NYC, Airbnb listings consist of 13,240 entire apartments and 5,181 private or shared rooms.

[18] See `http://pando.com/2013/12/08/airbnb-says-this-man-does-not-exist-so-i-had-coffee-with-him/` for the story of an Airbnb host who has received outside funding to build a business leasing out apartments on Airbnb.

inventory in Austin, their listings account for 45% of the total number of guest reviews, suggesting that these units serve greater demand than the units managed by single-unit hosts. Therefore, the expected hotel revenue gains of regulation that limits the number of listings per host may be greater than our simulations suggest.

7 Sensitivity Analyses and Robustness Checks

We now study the sensitivity of our results to modeling assumptions we made, and identify several potential concerns that could limit the applicability and generality of our results. We address these concerns through a series of robustness checks which result in estimations that continue to be consistent with our primary specification.

7.1 Endogeneity of Airbnb market entry

A key concern in studies like ours is endogeneity in adoption, *i.e.*, whether there are time-varying unobservables that are correlated both with Airbnb adoption and hotel revenues. For example, failure to control for unemployment rate would raise serious concerns, as clearly it is negatively correlated with hotel revenues, and it is natural to hypothesize that it could be positively correlated with Airbnb adoption. Indeed, Airbnb touts the help it provides to struggling homeowners in paying their mortgage, occasionally making a difference in these properties being repossessed by lenders.[19] While we include a rich set of controls in our analyses, and employ a difference-in-differences methodology, the issue potentially remains: are there city-specific, time-varying unobservables that affect Airbnb adoption as well as hotel revenues? For example, do individuals wait for hotel revenues in their vicinity to drop to list their property on Airbnb?

Intuitively, we view this as unlikely, given the distributed and idiosyncratic nature of Airbnb's expansion. Airbnb adoption is driven by a broad set of distributed individual hosts, each acting in their own self-interest, as opposed to a planned rollout from corporate headquarters. While a firm like Airbnb could promote itself through incentives when and where it views its competition from hotels as weak, we have no evidence that Airbnb engaged in this strategy. Instead, individuals, many with little prior experience in this marketplace, acted to list their inventory on Airbnb. Since the overhead of doing so is low (limited mostly to production costs in setting up a web page to market the property – there are no fees associated with creating a new Airbnb listing), we argue that individuals largely have acted

[19] See "How Airbnb helps users save their homes", August 2012, `http://finance.fortune.cnn.com/2012/08/16/Airbnb-foreclosure/`

opportunistically, rather than leveraging deeper strategic decision-making regarding hotel performance, when they adopt Airbnb. This line of reasoning suggests that reverse-causality is not an issue: it is unlikely that many individuals strategically consider hotel revenues in their vicinity before listing their property on Airbnb.

To provide additional quantitive evidence against such concerns, we analyze three different subsets of our dataset: 1) cities which Airbnb has entered by Q2 2013, 2) the top-10 largest cities in Texas, and 3) the top-5 largest cities in Texas. By focusing on progressively smaller sections of the Texas market, we limit the scope of self-selection in decisions to adopt Airbnb that are correlated with city-specific unobservables. We report the results of these robustness checks in Table 6. In all cases, the impact of Airbnb remains statistically significant, and quantitatively similar to our main specification.

To further rule out the possibility of endogeneity in Airbnb's market entry, we employ hazard models to predict the entry of Airbnb into cities as a function of demographic factors, other socioeconomic factors, and hotel revenue. Table 7 presents the results of this analysis. In column one we define Airbnb market entry as the year and quarter in which the first listing appears in the market. In columns two and three we use alternative definitions of Airbnb market entry: when at least 10 Airbnb listings appear in the market and the date of the market's first review respectively. We find that $\textit{log hotel rooms}_{it}$ (column one, two and three), $\textit{unemployment rate}_{it}$ (column two and three) and $\textit{log Accommodation Sector Employees}_{it}$ (column two) are statistically significant predictors of Airbnb's entry. However, we find no direct relationship between hotel revenue and Airbnb adoption. These results further support the argument that hotel revenues and the timing of Airbnb's market entry are unrelated.

7.2 Alternative definitions of Airbnb penetration

We next address the concern that our results are driven by our particular choice of penetration metric. In response to this concern, we perform sensitivity analyses with alternative measures. First, we consider Airbnb-entry as a dummy variable which is zero before the *first* Airbnb listing appears in a city, and one thereafter. Our results are reported in the first column of Table 8. The coefficient for Airbnb-entry is near zero. This entry measure appears in the main specification of other studies including Forman et al. (2009), and Chan and Ghose (2011). But is it equally valid in the Airbnb setting? According to this definition, and referring back to Table 2, Airbnb has penetrated the San Antonio market in an identical manner to the Austin market. However, by 2013 these similarly sized cities have seen wildly different levels of Airbnb adoption, weakening the case for the appropriateness of this metric.

Motivated by this observation, we experiment with two more robust definitions of market entry: 1) when at least 10 Airbnb listings appear in a city, and 2) when at least 50 Airbnb listings appear in a city. We define two corresponding dummy entry variables, and report our results in columns two and three of Table 8. In both instances, the coefficient for Airbnb entry is negative, statistically significant, and of substantive magnitude.

Next, based on the observation that Airbnb only allows guests who have paid for accommodation to review a listing, we construct a measure of Airbnb penetration based on the count of *reviewed* listings in each city. First, we exclude any listing which has never been reviewed. Then, for the entry date of each remaining listing, we use the date of its first review (rather than the date the listing owner registered on Airbnb, used in our preferred specification). This measure clearly reflects a more conservative estimate of Airbnb's diffusion, as the date of the first review occurs strictly later than the appearance of the property on Airbnb. Our results using this conservative measure are reported in the fourth column of Table 8 and show a negative, and statistically significant relationship, albeit with smaller magnitude than our preferred specification. Taken together, these robustness checks suggest that our results are not driven by our particular selection of penetration measure.

7.3 The 2008 financial crisis

Our next concern regards the 2008 financial crisis, a complex economic event whose impacts are still being tallied. Airbnb entered the Texas market during the 3rd quarter of 2008, concurrent with Lehman's bankruptcy filing and Merrill Lynch's acquisition by Bank of America. Recall further that the crisis was driven in no small part by subprime lending, which is directly related to Airbnb, a business driven by spare housing inventory. The impact of the crisis on hotel revenues is most evident in the right panel of Figure 8. Given the complexity of the event, one has to wonder if we can control for every time-varying unobservable that could affect revenues as well as Airbnb adoption. Our approach is simple, but not especially nuanced: we exclude the period from the third quarter of 2008 through the end of 2009, which spans the main crisis period. The results using this censored dataset are reported in Table 9, with the full estimation across different hotel price segments. Overall, our results remain unchanged. While somewhat surprising, we conjecture that between controlling for time fixed effects and city-specific time trends we capture the "essence" of the crisis in our prior specification.

7.4 Alternative levels of geographic aggregation

A natural question that arises from our analysis regards the level of geographic aggregation over which we define each competitive market. Recall that our preferred specification used city-level granularity, the relevance of which we argued for in Section 4. Another alternative, which we now consider, is to use the finer ZIP code level granularity in defining markets.

Defining markets using ZIP codes has some limitations. Each hotel's placement within a city is typically the result of strategic placement of a firm, and hotels are therefore often concentrated in ZIP codes zoned specifically for development. See for example the highly non-uniform placement of hotels in Travis County, TX, in the bottom-right panel of Figure 4. In contrast, Airbnb hosts enter the market opportunistically and presumably pay little attention to their ZIP code in their decision to enter the market. Indeed, the histograms in Figure 7 confirm that Airbnb and hotel frequencies do not respect ZIP code boundaries in Austin. Here, we plot the number of hotels and Airbnb listings for the ten Austin ZIP codes with the most hotels (left panel) and with the most Airbnb listings (right panel) as of Q1 2013.

Despite this limitation, ZIP code level aggregation provides us with a robustness check against mistakingly associating unobserved, city-level tourism trends, with an Airbnb-driven effect. We ran regressions using this alternative definition by taking the treatment variable in our analysis to be the log of the number of Airbnb listings in the ZIP code of (as opposed to the city of) hotel i at time t. We report the results for the DDD models previously presented in Section 5.1 with ZIP code level aggregation in Table 10. The nature of our findings remains unchanged with respect to the results for city-level granularity. These results suggest that the Airbnb effect on hotel revenues is not driven by our particular choice of geographic aggregation.

8 Discussion and Future Work

The sharing economy has recently emerged as a viable alternative to fulfilling a variety of consumer needs, ranging from prepared meals to cars to overnight accommodations, that were previously provided by firms. As the size of the sharing economy has grown, so has the magnitude of its economic and societal impacts. To date, discussion of the sharing economy has risen to the level of heated debate, along with a rapidly growing level of attention by legal and regulatory bodies. But rigorous studies that attempt to empirically estimate the impacts of the sharing economy have not yet emerged. Our work is among the first to provide empirical evidence that the sharing economy is significantly *changing* consumption patterns, as opposed to generating purely incremental economic activity, as argued in prior work. Studying the case of Airbnb, a pioneer in shared accommodations, we identify that its entry into the Texas market has had a quantifiable negative impact on local hotel revenues. The substitution patterns we observe by incorporating business hotels and high-end hotels as control groups, strongly suggest that Airbnb provides a viable, but imperfect, alternative for certain traditional types of overnight accommodation. Our results, which we confirmed through a series of robustness checks, pinpoint lower-end hotels as those that are most vulnerable to increased competition from rentals enabled by firms like Airbnb. Using counterfactual simulations, we also model various regulatory interventions to provide an intuitive quantification of their likely effects on Airbnb and hotel revenue in Texas.

Returning to the thesis that the sharing economy has the potential to transformatively increase social welfare, as evangelized by Botsman (2012) and others, we assert that a large population of individuals worldwide have indeed benefitted from Airbnb: both those that derive incremental income by renting properties through Airbnb, as well as individuals selecting an Airbnb rental as an alternative to a hotel stay. But more broadly, our results should be viewed from outside the confines of the accommodation industry. This more encompassing viewpoint can weigh the positive change the sharing economy can bring about not only by providing imperfect substitutes for existing products, but also, through an application of Say's Law, by generating demand that did not previously exist through the supply of new products and services. Harkening back to arguments Airbnb has made, supply of inexpensive accommodations can increase travel and tourism spend overall, and thus the sharing economy could be a net producer of new jobs. However, these positives must be evaluated against various costs, including those estimated in this paper. Our study represents a first step into understanding the complex economic, regulatory, and technological issues surrounding the sharing economy. With the projected rapid growth of the sharing economy, a host of related studies will be needed to fully understand and reap its benefits.

References

Baker, Kathleen, Alex Coulter. 2007. Terrorism and Tourism: The Vulnerability of Beach Vendors' Livelihoods in Bali. *Journal of Sustainable Tourism* **15**(3) 249–266.

Bertrand, Marianne, Esther Duflo, Sendhil Mullainathan. 2004. How much should we trust differences-in-differences estimates? *The Quarterly Journal of Economics* **119**(1) 249–275.

Botsman, Rachel. 2012. The currency of the new economy is trust. [Video file]. `http://www.ted.com/talks/rachel_botsman_the_currency_of_the_new_economy_is_trust.html`. [Online].

Brynjolfsson, Erik, Yu Jeffrey Hu, Mohammad S. Rahman. 2009. Battle of the Retail Channels: How Product Selection and Geography Drive Cross-Channel Competition. *Management Science* **55**(11) 1755–1765.

Cervero, Robert, Aaron Golub, Brendan Nee. 2007. City CarShare: Longer-Term Travel Demand and Car Ownership Impacts. *Transportation Research Record: Journal of the Transportation Research Board* **1992**(1) 70–80.

Chan, Jason, Anindya Ghose. 2011. Internet's Dirty Secret: Assessing the Impact of Technology Shocks on the Outbreaks of Sexually Transmitted Diseases. *Workshop of Health IT and Economics (WHITE)* .

Forman, Chris, Anindya Ghose, Avi Goldfarb. 2009. Competition Between Local and Electronic Markets: How the Benefit of Buying Online Depends on Where You Live. *Management Science* **55**(1) 47–57.

Goldfarb, Avi, Catherine Tucker. 2011a. Advertising Bans and the Substitutability of Online and Offline Advertising. *Journal of Marketing Research* **48**(2) 207–227.

Goldfarb, Avi, Catherine Tucker. 2011b. Search Engine Advertising: Channel Substitution When Pricing Ads to Context. *Management Science* **57**(3) 458–470.

Goldfarb, Avi, Catherine Tucker. 2011c. Substitution Between Offline and Online Advertising Markets. *Journal of Competition Law and Economics* **7**(1) 37–44.

King, Gary, Michael Tomz, Jason Wittenberg. 2000. Making the most of statistical analyses: Improving interpretation and presentation. *American Journal of Political Science* 347–361.

Kosova, Renata, Cathy A. Enz. 2012. Terrorism and The Financial Crisis of 2008: The Impact of External Shocks on US Hotel Performance. *Cornell Hospitality Quarterly* .

Kroft, Kory, Devin G. Pope. 2013. Does Online Search Crowd Out Traditional Search and Improve Matching Efficiency? Evidence from Craigslist. *Journal of Labor Economics* (Forthcoming).

Martin, Elliot, Susan A. Shaheen, Jeffrey Lidicker. 2010. Impact of carsharing on household vehicle holdings. *Transportation Research Record: Journal of the Transportation Research Board* **2143**(1) 150–158.

Mundlak, Yair. 1978. On the pooling of time series and cross section data. *Econometrica: journal of the Econometric Society* 69–85.

O'Connor, Noëlle, Mary Rose Stafford, Gerry Gallagher. 2008. The Impact of Global Terrorism on Ireland's Tourism Industry: an Industry Perspective. *Tourism and Hospitality Research* **8**(4) 351–363.

Seamans, Robert, Feng Zhu. 2013. Responses to Entry in Multi-Sided Markets: The Impact of Craigslist on Local Newspapers. *Management Science* (Forthcoming).

Wolfers, Justin. 2006. Did Unilateral Divorce Laws Raise Divorce Rates? A Reconciliation and New Results. *The American Economic Review* **96**(5) 1802–1820.

Table 1: Means and standard deviations (in parentheses) of the variables used in our analysis.

Year	Airbnb Listings	Hotel Revenue ($)	Unemployment Rate	Hotel Rooms	Population	Accommodation Sector Employees
2003	0.0	268392.5	6.6	13214.2	1152272.3	129.3
	(0.0)	(583259.5)	(1.3)	(17302.1)	(1260860.8)	(96.9)
2004	0.0	283580.3	6.0	13594.3	1187167.6	134.6
	(0.0)	(648198.6)	(1.2)	(17800.1)	(1290309.2)	(99.6)
2005	0.0	303021.7	5.3	13364.4	1203840.7	138.0
	(0.0)	(717413.1)	(1.0)	(17556.6)	(1308677.9)	(101.8)
2006	0.0	333977.2	4.9	13377.7	1228555.9	142.1
	(0.0)	(781584.8)	(0.9)	(17483.0)	(1345337.5)	(105.8)
2007	0.0	339360.9	4.3	13298.3	1233473.5	147.6
	(0.0)	(765206.8)	(0.8)	(17623.4)	(1359674.2)	(110.7)
2008	0.6	337071.9	4.9	13217.0	1241844.1	150.9
	(3.0)	(748809.8)	(1.0)	(17546.6)	(1377469.2)	(113.2)
2009	4.1	270484.0	7.4	13055.0	1245851.2	149.9
	(16.4)	(592641.9)	(1.4)	(17531.9)	(1401038.8)	(112.5)
2010	13.3	269551.5	8.1	13626.2	1261781.9	151.2
	(51.3)	(603565.4)	(1.4)	(18454.4)	(1418697.4)	(112.2)
2011	55.5	284115.4	7.8	13675.9	1283574.3	157.2
	(211.1)	(638167.1)	(1.6)	(18653.9)	(1450474.8)	(115.6)
2012	160.0	301599.9	6.7	14106.1	1309857.6	164.9
	(634.7)	(652575.6)	(1.5)	(19219.3)	(1478980.8)	(121.4)
2013	257.9	333018.3	6.5	14358.2	1332087.0	170.2
	(993.5)	(736259.8)	(1.5)	(19646.9)	(1506317.7)	(125.4)
Overall	45.7	300622.8	6.4	13545.4	1250082.9	150.2
	(355.1)	(677674.0)	(1.8)	(18153.4)	(1394690.3)	(112.3)
N	118687	118687	118576	118687	118437	97836

Note: 2013 statistics only include Q1 and Q2.

Table 2: Airbnb's spatial and temporal penetration. Cumulative counts of Airbnb listings per year in the ten most populous Texas cities.

(Pop.)	Houston 2.16M	San Antonio 1.38M	Dallas 1.24M	Austin 0.84M	Ft. Worth 0.78M	El Paso 0.67M	Arlington 0.38M	Corpus Christi 0.31M	Plano 0.27M	Laredo 0.24M
2008	1	5	0	23	0	0	0	0	0	0
2009	4	8	7	125	2	0	0	0	0	0
2010	30	15	19	375	7	0	2	0	0	0
2011	129	52	68	1464	23	3	9	6	3	1
2012	254	127	150	4278	44	7	17	21	11	1
2013	353	155	198	5468	52	15	20	25	16	1

Table 3: The impact of Airbnb on hotel revenue.

	(1)	(2)	(3)	(4)	(5)
log Airbnb Listings	−0.040***	−0.037***	−0.041***	−0.041***	−0.050***
	(−8.37)	(−7.74)	(−8.51)	(−8.60)	(−9.99)
Unemployment Rate		−0.052***	−0.052***	−0.052***	−0.049***
		(−9.86)	(−9.95)	(−10.06)	(−7.44)
log Hotel Rooms			−0.284***	−0.286***	−0.289***
			(−7.87)	(−7.88)	(−6.04)
log Population				−0.158	−0.270**
				(−1.44)	(−2.11)
log Accommodation Sector Employees					2.028***
					(14.32)
N	118687	118576	118576	118437	97836
R^2 within	0.20	0.20	0.20	0.20	0.19

Note: The dependent variable is log Hotel Revenue$_{it}$. Cluster-robust t-statistics (at the individual hotel level) are shown in parentheses. All specifications include hotel and time fixed effects, and a city-specific quadratic time trend.
Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

Table 4: The impact of Airbnb on hotel revenue by hotel price segment.

	Budget	Economy	Midprice	Upscale	Luxury	DDD
log Airbnb Listings	−0.066***	−0.047***	−0.040***	−0.039***	−0.021	−0.025***
	(−7.05)	(−4.29)	(−4.05)	(−3.58)	(−1.29)	(−4.21)
Unemployment Rate	−0.052***	−0.023*	−0.045***	−0.080***	−0.001	−0.049***
	(−4.55)	(−1.65)	(−3.84)	(−5.40)	(−0.04)	(−7.40)
log Hotel Rooms	−0.169***	−0.314***	−0.417***	−0.567***	−0.010	−0.289***
	(−3.92)	(−6.02)	(−8.12)	(−7.43)	(−0.12)	(−6.04)
log Population	−0.456**	−0.519**	0.311	−0.052	−0.603	−0.239*
	(−2.24)	(−2.25)	(1.18)	(−0.17)	(−1.58)	(−1.90)
log Accommodation Sector Employees	1.601***	3.289***	2.292***	1.604***	3.098*	2.030***
	(9.99)	(8.68)	(6.64)	(4.38)	(1.67)	(14.32)
Budget × log Airbnb Listings						−0.037***
						(6.56)
Economy × log Airbnb Listings						−0.044***
						(−8.15)
Midprice × log Airbnb Listings						−0.015***
						(−2.89)
Upscale × log Airbnb Listings						−0.004
						(−0.89)
N	37387	20816	19179	12808	7646	97836
R^2 within	0.22	0.30	0.24	0.21	0.16	0.19

Note: The dependent variable is log Hotel Revenue$_{it}$. Cluster-robust t-statistics (at the individual hotel level) are shown in parentheses. All specifications include hotel and time fixed effects, and a city-specific quadratic time trend.
Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

Table 5: The impact of Airbnb on hotel revenues by meeting space

	No Meeting Space	Meeting Space at least 500 sq. ft.	Meeting Space at least 1000 sq. ft.	Meeting Space at least 2000 sq. ft.	DDD
log Airbnb Listings	−0.055***	−0.041***	−0.032***	−0.031**	−0.057***
	(−7.46)	(−4.98)	(−3.12)	(−2.33)	(−10.79)
Unemployment Rate	−0.047***	−0.054***	−0.053***	−0.060***	−0.049***
	(−4.91)	(−5.12)	(−3.66)	(−3.15)	(−7.40)
log Hotel Rooms	−0.236***	−0.273**	−0.176	−0.121	−0.289***
	(−5.96)	(−2.41)	(−1.46)	(−0.96)	(−6.04)
log Population	−0.397***	0.162	0.032	0.018	−0.260**
	(−2.68)	(0.75)	(0.13)	(0.05)	(−2.03)
log Accommodation Sector Employees	2.142***	1.740***	1.684***	1.668***	2.030***
	(11.85)	(5.87)	(4.29)	(4.17)	(14.32)
log Meeting Space × log Airbnb Listings					0.002***
					(3.82)
N	51532	33692	21412	12734	97836
R^2 within	0.22	0.19	0.20	0.22	0.19

Note: The dependent variable is log Hotel Revenue$_{it}$. Cluster-robust t-statistics (at the individual hotel level) are shown in parentheses. All specifications include hotel and time fixed effects, and a city-specific quadratic time trend.

Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

Table 6: Robustness checks for endogeneity in Airbnb's entry into different cities.

	Airbnb Eventually	Top 10 Cities	Top 5 Cities
log Airbnb Listings	−0.053***	−0.083***	−0.076***
	(−10.57)	(−12.17)	(−8.88)
Unemployment rate	−0.034***	−0.032***	−0.006
	(−4.77)	(−2.78)	(−0.44)
log Hotel Rooms	−0.337***	−0.419***	0.020
	(−9.06)	(−5.68)	(0.21)
log Population	−0.216	−0.247	−0.479
	(−1.51)	(−0.99)	(−1.56)
log Accommodation Sector Employees	2.107***	2.212***	2.728***
	(13.29)	(10.80)	(12.68)
N	84489	47392	38109
R^2 within	0.16	0.14	0.14

Note: The dependent variable is log Hotel Revenue$_{it}$. Cluster-robust t-statistics (at the individual hotel level) are shown in parentheses. All specifications include hotel and time fixed effects, and a city-specific quadratic time trend.

Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

Table 7: Hazard Models for predicting Airbnb's entry

	Airbnb enters when 1 listing appears	Airbnb enters when 10 listings appear	Airbnb enters when 1 listing is reviewed
log Population	0.048	−0.190	0.041
	(0.62)	(−0.88)	(0.37)
log Hotel Rooms	0.551***	1.286***	0.622***
	(6.49)	(4.28)	(4.90)
Unemployement Rate	−0.102	−0.280**	−0.189*
	(−1.40)	(−2.10)	(−1.82)
log Accommodation Sector Employees	0.103	0.307*	0.140
	(1.54)	(1.96)	(1.50)
log Hotel Revenue	0.146	0.043	0.071
	(1.17)	(0.14)	(0.41)
N	9892	10607	10408
Log Likelihood	-593.6	-97.7	-361.2

Note: The table reports the results from the hazard model predicting Airbnb's entry into a city. The dependent variable equals one when the city experiences Airbnb's entry and zero otherwise. Airbnb's entry is an absorbing state, therefore the city is dropped from the sample in the quarters after the dependent variable becomes one. The controls include demographic and number of employees in the accommodation sector at county level and unemployment rate, number of hotel rooms and hotel revenue at city level.

Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

Table 8: Robustness checks with alternate definitions of Airbnb penetration.

	Airbnb enters when 1 listing appears	Airbnb enters when 10 listings appear	Airbnb enters when 50 listings appear	Airbnb enters when 1 listing is reviewed
Airbnb Entry	0.003	−0.072***	−0.066***	−0.017**
	(0.44)	(−8.44)	(−6.39)	(−2.42)
Unemployment Rate	−0.051***	−0.049***	−0.052***	−0.051***
	(−7.72)	(−7.37)	(−7.80)	(−7.69)
log Hotel Rooms	−0.276***	−0.282***	−0.276***	−0.277***
	(−5.95)	(−5.96)	(−5.96)	(−5.97)
log Population	−0.269**	−0.264**	−0.278**	−0.265**
	(−2.10)	(−2.06)	(−2.18)	(−2.07)
log Accommodation Sector Employees	1.992***	2.048***	1.981***	1.992***
	(14.12)	(14.42)	(14.06)	(14.12)
N	97836	97836	97836	97836
R^2 within	0.18	0.18	0.18	0.18

Note: The dependent variable is log Hotel Revenue$_{it}$. Cluster-robust t-statistics (at the individual hotel level) are shown in parentheses. All specifications include hotel and time fixed effects, and a city-specific quadratic time trend.

Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

Table 9: The impact of Airbnb on hotel revenues excluding observations spanning the 2008 financial crisis.

	All	Budget	Economy	Midprice	Upscale	Luxury
log Airbnb	−0.046***	−0.065***	−0.032**	−0.044***	−0.040***	−0.012
Listings	(−8.38)	(−6.19)	(−2.49)	(−4.28)	(−3.35)	(−0.63)
Unemployment	−0.023***	−0.036***	0.005	−0.014	−0.045***	0.034
Rate	(−3.06)	(−2.73)	(0.32)	(−1.05)	(−2.65)	(1.13)
log Hotel Rooms	−0.315***	−0.199***	−0.308***	−0.413***	−0.604***	−0.007
	(−7.82)	(−4.99)	(−6.79)	(−8.12)	(−7.34)	(−0.08)
log Population	−0.120	−0.562***	−0.286	0.684**	0.348	−0.698*
	(−0.86)	(−2.73)	(−1.19)	(2.51)	(0.99)	(−1.80)
log Accommodation	2.045***	1.658***	3.253***	2.364***	1.517***	3.315*
Sector Employees	(14.38)	(10.36)	(9.14)	(6.61)	(4.01)	(1.77)
N	82823	31803	17625	16198	10751	6446
R^2 within	0.20	0.23	0.31	0.26	0.22	0.18

Note: The dependent variable is log Hotel Revenue$_{it}$. Cluster-robust t-statistics (at the individual hotel level) are shown in parentheses. All specifications include hotel and time fixed effects, and a city-specific quadratic time trend.

Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

Table 10: DDD when the relevant market considered is a ZIP code.

	DDD Price Category	DDD Meeting Space
log Airbnb Listings (ZIP Code)	0.011	−0.024***
	(0.92)	(−2.78)
Unemployment Rate	−0.050***	−0.050***
	(−7.37)	(−7.38)
log Hotel Rooms (ZIP Code)	−0.160***	−0.160***
	(−8.18)	(−8.20)
log Population	−0.132	−0.140
	(−1.17)	(−1.23)
log Accommodation Sector Employees	1.986***	1.984***
	(14.04)	(14.03)
log Meeting Space × log Airbnb Listings (ZIP Code)		0.003***
		(2.82)
Budget × log Airbnb Listings (ZIP Code)	−0.044***	
	(−2.82)	
Economy × log Airbnb Listings (ZIP Code)	−0.060***	
	(−3.85)	
Midprice × log Airbnb Listings (ZIP Code)	−0.005	
	(−0.44)	
Upscale × log Airbnb Listings (ZIP Code)	0.010	
	(0.90)	
N	97836	97836
R^2 within	0.23	0.23

Note: The dependent variable is log Hotel Revenue$_{it}$. Cluster-robust t-statistics (at the individual hotel level) are shown in parentheses. All specifications include hotel and time fixed effects, and a ZIP code-specific quadratic time trend.

Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

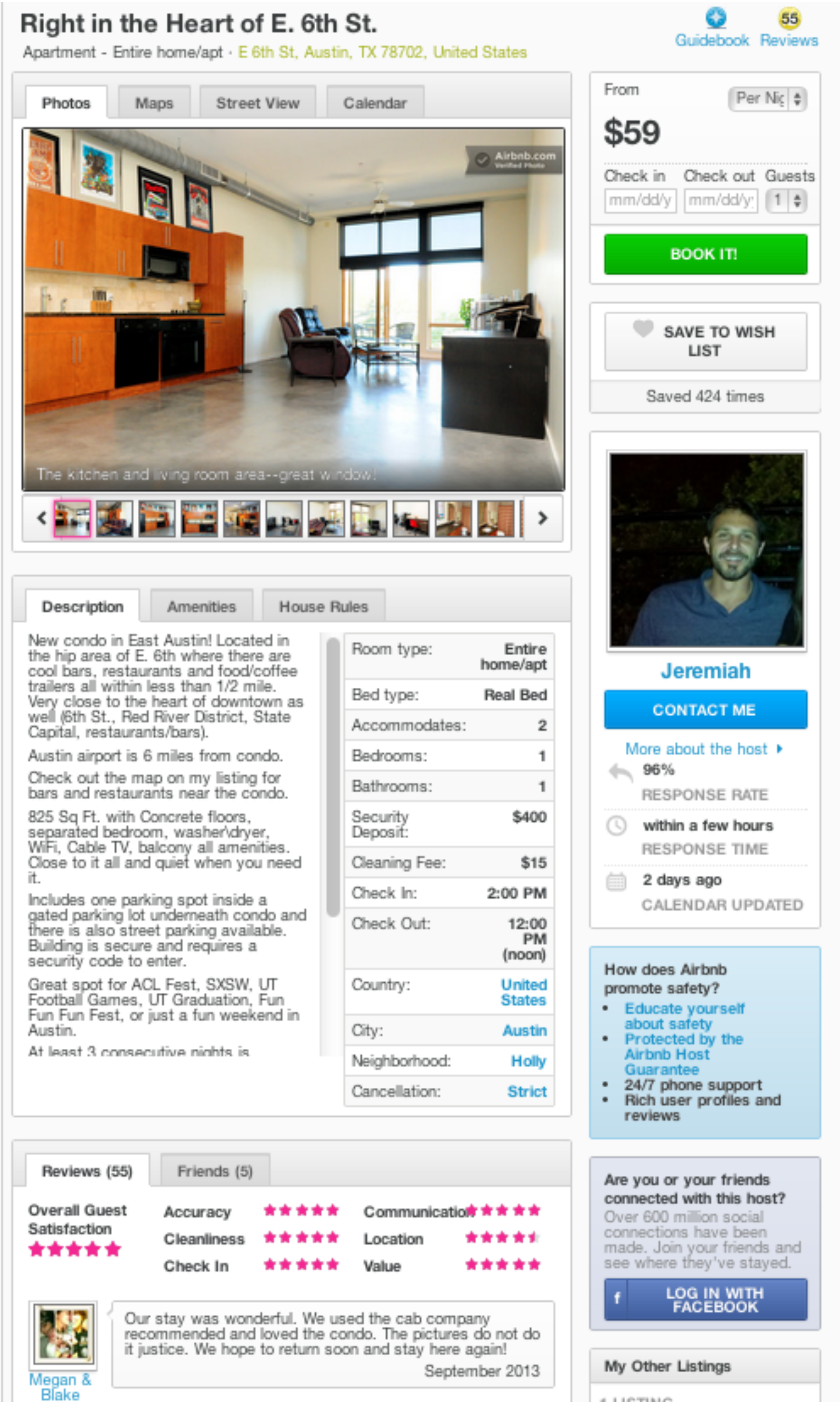

Figure 1: A typical Airbnb listing.

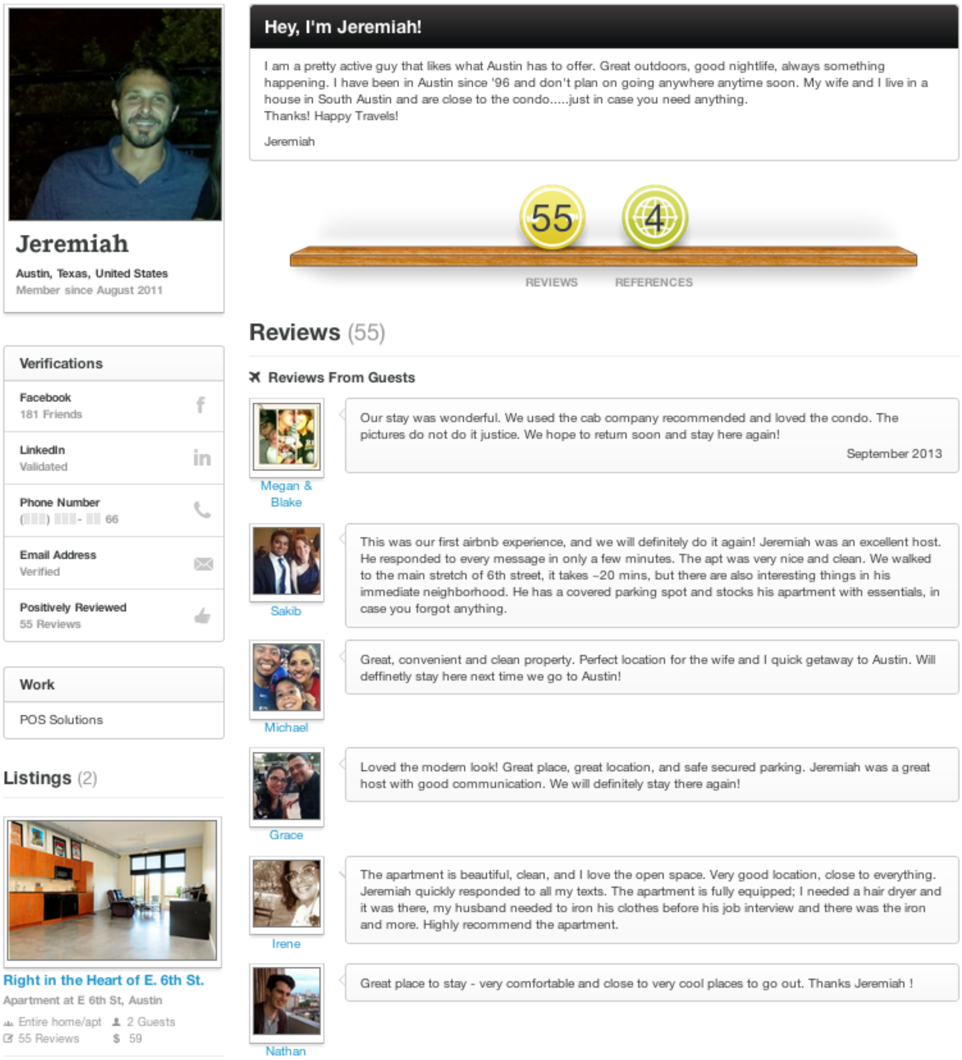

Figure 2: A typical Airbnb user profile.

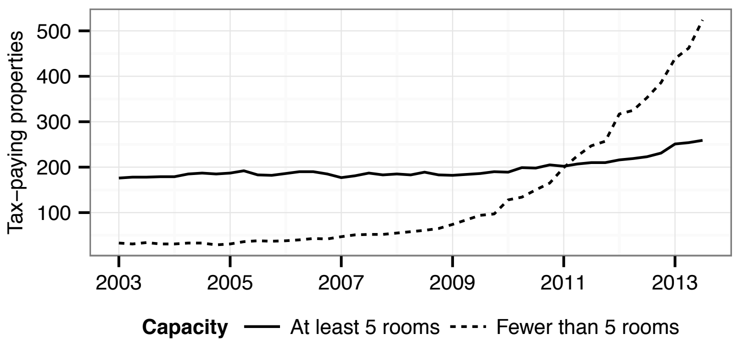

Figure 3: Quarterly counts of Austin hotel occupancy tax paying properties by capacity.

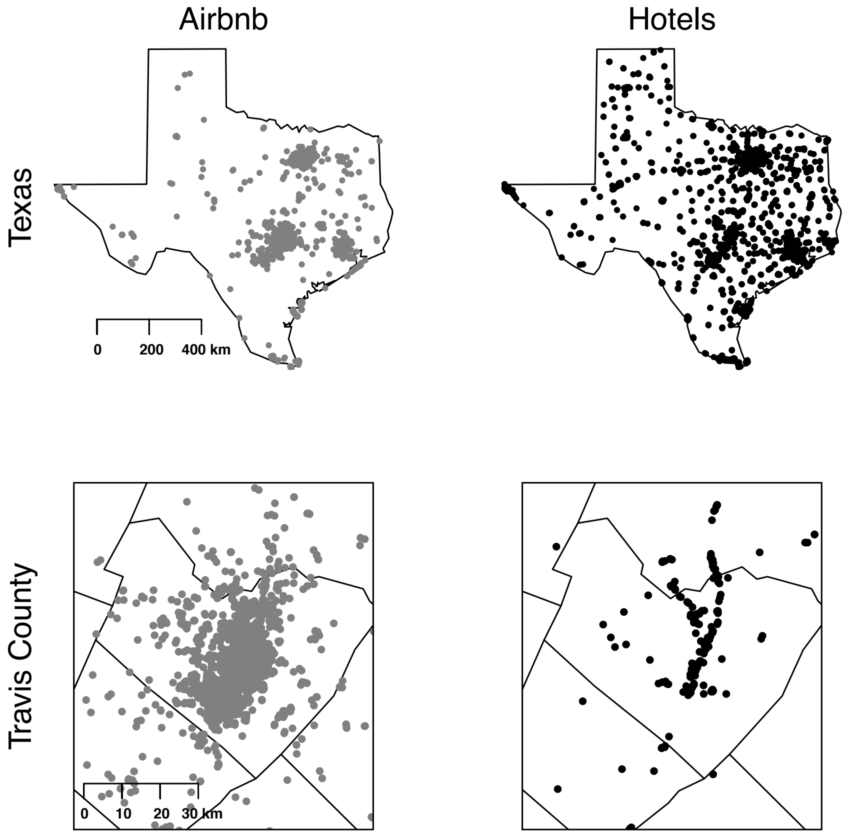

Figure 4: Geographical distribution of hotels and Airbnb listings in the state of Texas (top) and in Travis County, TX (bottom) in 2013.

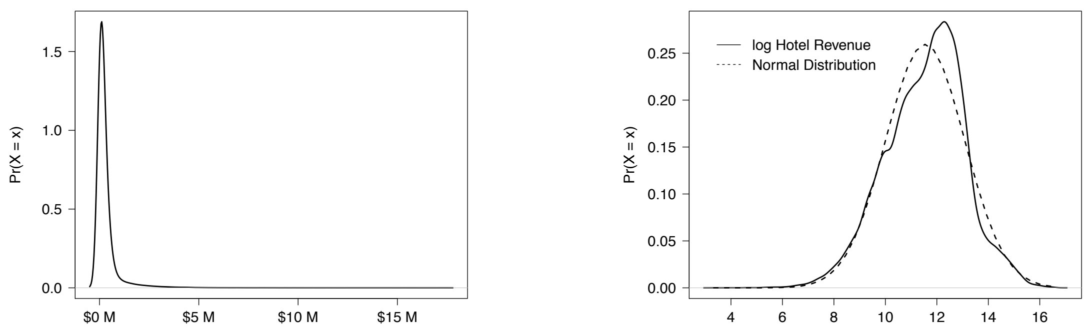

Figure 5: Density of *Hotel Revenue* (left) and *log Hotel Revenue* (right).

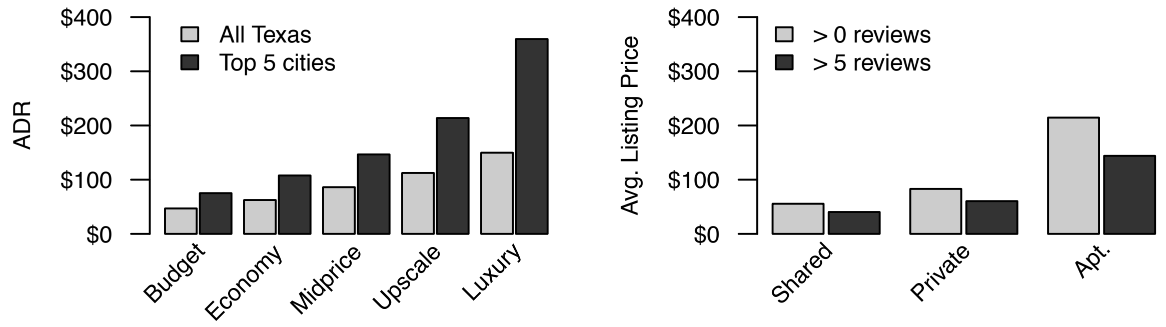

Figure 6: Average 2012 daily rate (ADR) by price segment (left) vs. average Airbnb listing price by type (right). In the left panel, light bars reflect ADR across all of Texas (source: STR), and dark bars reflect ADR for the five biggest metro areas of Texas (source: Hotels.com). The right panel presents average price by Airbnb listing type in Texas in 2013.

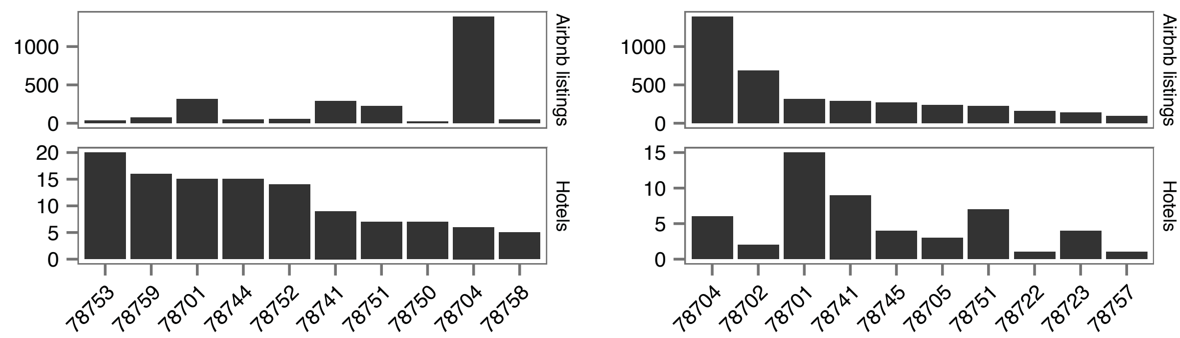

Figure 7: Hotel and Airbnb density, as measured in Q2 2013, for the ten Austin zip codes with the most hotels (left) and with the most Airbnb listings (right).

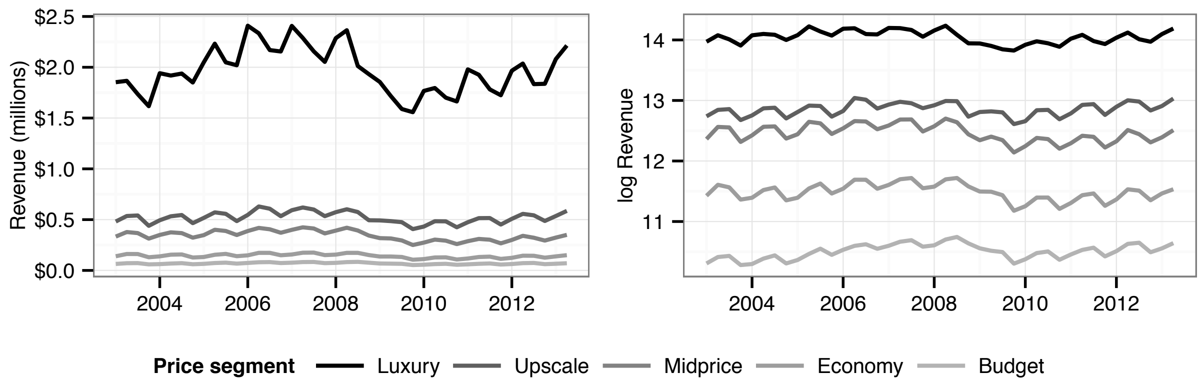

Figure 8: Average quarterly hotel revenue by price segment (left) and log hotel revenue (right). Dollar values are deflated to January 2003 (start of our sample).

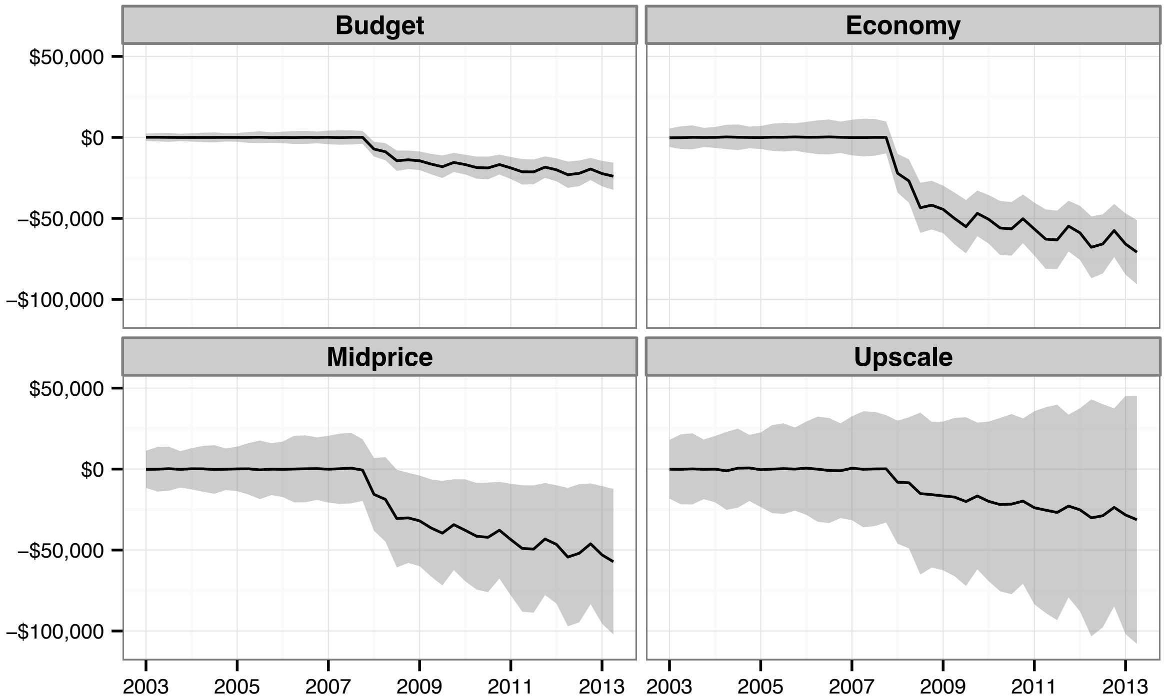

Figure 9: Expected difference in the quarterly revenues of the average Austin hotel between the cases of a) observed Airbnb penetration, and b) 1% of observed Airbnb penetration.

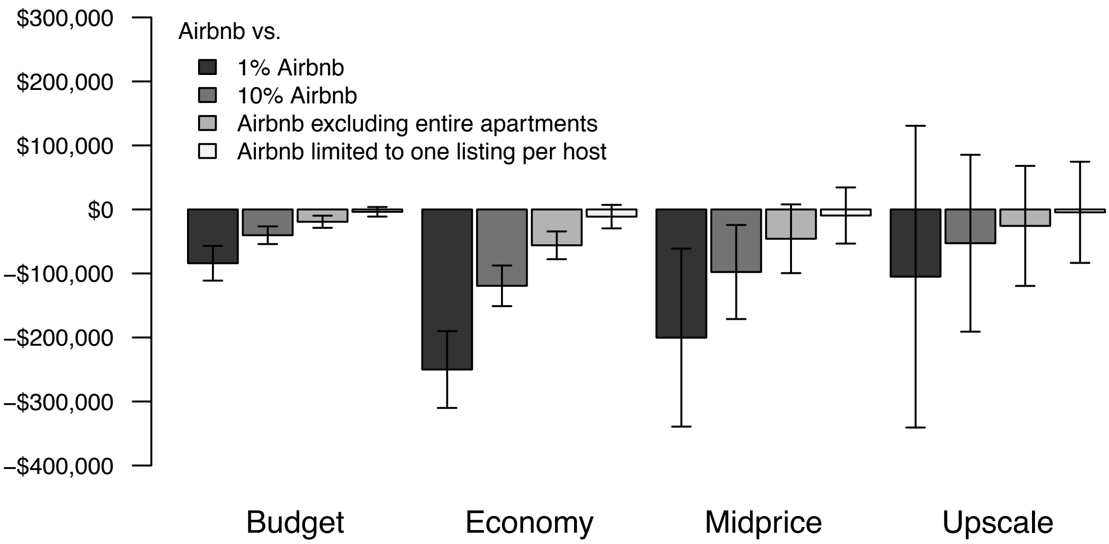

Figure 10: Expected difference in 2012 revenues for the average Austin hotel between a) observed Airbnb penetration, and b) four counterfactual scenarios with varying levels of Airbnb penetration.