Airbnb's Role in Tourism Gentrification

Working Paper

Jake Schild*

September 22, 2018

Abstract

As Airbnb has rose in popularity so to has the literature's interest in understanding the economic impact of this new technology. Up to this point the literature has largely focused on the disruptive effects Airbnb has had on the hotel industry and housing market. However, the proliferation of Airbnb has also resulted in the redistribution of tourists within cities which has the potential to impact the development of businesses in the entertainment and service sectors within a city neighborhood. As businesses respond to the changing demographics of the area, an increase in tourism, and therefore an increase in demand for entertainment and consumption, an neighborhood can quickly become an agglomeration of bars, restaurants, and other tourist attractions, more commonly known within the literature as tourism gentrification. The goal of this paper is to analyze the relationship between the intensity of Airbnb usage with a neighborhood and local business development. Two conjectures are drawn from a theoretical model based on a Krugman style model of trade, and then empirically tested using a collection of novel data sources. The results suggest an increase in Airbnb usage leads to an increase in the number of entertainment sector firms within the area, but has no statistically significant effect on the number of firms in the service sector.

JEL Code: L89, R12, Z32
Key Words: Airbnb, Tourism Gentrification

*Indiana University Email: jjschild@indiana.edu

1 Introduction

Airbnb was first introduced in 2007 as a solution to two problems. Founders Joe Gebbia and Brian Chesky were looking for a way to help pay their rent, and there was a shortage of accommodations for attendants of a San Francisco design conference.[1] Since then Airbnb has spread to more than 190 countries, has had over 300 million guests, and has become one of the largest lodging companies in the world.[2] Its rapid rise in popularity and lack of regulation has led to Airbnb being viewed among researchers as a disruptive innovation within the hotel industry (Guttentag and Smith, 2017; Zervas et al., 2017) and housing market (Lee, 2016; Schäfer and Braun, 2016; Gurran and Phibbs, 2017). What's more, cities including San Francisco (Brousseau et al., 2015) and New York (New York State Attorney General, 2014) have conducted their own studies documenting the effect Airbnb has had on rental availability and costs of affordable housing. However Airbnb has also had unintended consequences on both local residents and tourism.

Because Airbnb utilizes residential accommodations it has changed the flow of tourism within a city. Prior to Airbnb tourists were largely limited to the hotels within a city's main commercial district, but now tourists are able to stay in more residential districts. Areas that have historically had little or no tourism. Unlike the impact on the hotel industry and housing market, minimal research has been conducted trying to understand the impact Airbnb and the resulting redistribution of tourists has had on residents and residential areas.

Of the literature analyzing the impact most have focused on Airbnb's benefits to the residents. A report published by Airbnb claims hosts earn on average $7,500 in additional income.[3] Furthermore, work by Alyakoob and Rahman (2018) shows increasing the intensity of Airbnb activity can lead to an increase in employment within the restaurant industry. Though this research documents some of the benefits of Airbnb it also inadvertently highlights one of the potential drawbacks. As more tourists utilize Airbnb and occupy a residential area the aggregate demand for goods and services within the area will change. For example, more tourists in the area implies a higher demand for restaurants which will lead to an increase in employment within the restaurant industry as documented by Alyakoob and Rahman. But it could also lead to a general increase in the number of restaurants and entertainment firms, forcing out other types of firms used primarily by residents. The goal of this paper is to analyze the relationship between the intensity of tourism with an area and local business development.

With the rise of urban tourism[4] in the latter half of the 20th century cities have put more emphasis on tourism as a source of revenue.[5] Specifically, they have invested substantial funds in new infrastructure, refurbishing, and developing new brand images (Judd, 1991). Any new infrastructure is typically centered around already established major tourist attractions (e.g. sports

[1]See *The Airbnb Story: How Three Ordinary Guys Disrupted an Industry, Made Billions... and Created Plenty of Controversy* for more details about the founding and subsequent rise of Airbnb.

[2]*Source:* `https://press.atairbnb.com/fast-facts/`

[3]*Source:* `https://wwwhttps://www.ftc.gov/system/files/documents/public_comments/2015/05/01740-96152.pdf`

[4]According to the World Trade Organization (UNWTO) urban tourism is defined as "...trips taken by travelers to cities or places of high population density. The duration of these trips is usually short (one to three days)." *Source:* `http://affiliatemembers.unwto.org/publication/global-report-city-tourism`

[5]Data from the UNWTO shows tourist arrivals have shown virtually uninterrupted growth since 1980. Furthermore, in their 2018 report the World Travel and Tourism Council states within the US tourism directly accounts for 2.6% of total GDP, and is forecasted to rise by 3.4% in 2018.

stadiums, convention centers), and has resulted in the creation of tourist districts, which usually correspond to a city's downtown.[6] However, the creation and proliferation of tourist districts does not go unmanaged. In his research, Getz (1993) discusses planning strategies for the creation of tourist districts, mentioning zoning as a means to encourage the desired development within a controlled area.

In order to promote the growth of the downtown area as a tourist destination it is zoned primarily for commercial use.[7] Conversely areas outside the downtown area are zoned for residential and limited commercial use. Consequently most tourist accommodations, like hotels, locate downtown which constricts tourists to residing downtown. Alternatively most residents choose to reside outside the downtown area in residential districts or neighborhoods. An unintended, or maybe intentional, consequence of zoning has been the separation of tourists and residents.

By separating tourists and residents the city is inadvertently isolating consumers by their preferences. Tourists have a strong preference for entertainment goods compared to residents who have a more diverse set of preferences, including a preference for residential goods and services. Both types of consumers can and do have at least some preference for goods of the other type, but have a stronger preference for the corresponding good. As a result, the local demand for entertainment firms is higher downtown, and the local demand for residential goods and services is higher in residential areas. Thus entertainment firms tend to agglomerate downtown while residential service firms cluster in residential districts. This result is equivalent to the Home Market Effect shown by **?**.

In his paper Krugman develops a model of international trade showing countries with identical tastes, technology, and factor endowments are still able to benefit from trade. He then uses this model to characterize the Home Market Effect, which is defined as the tendency for countries to export those goods in which they have a larger domestic market. Though it is not a typical application, the model Krugman can be adapted to model the "trade" that occurs within a city.[8] The districts or neighborhoods within a city can be interpreted as the regions or countries within the model. Each region is endowed with a population of consumers who have heterogeneous preferences. Firms then decide where to locate based on the size of the "domestic" demand within the region, and "trade" occurs when consumers travel to consume the products offered by firms in a different region. According to the model when tourists and residents are sorted into the downtown and residential districts, respectively, entertainment firms locate downtown and residential firms locate in the residential district. However, if consumers are redistributed such that domestic demand within a region changes firms will choose to relocate.

One consequence of Airbnb is the redistribution of tourists from the downtown area into residential districts.[9] As the number of tourists within a residential district increases so too does the local demand for entertainment. Thus more entertainment firms should choose to locate in the area.

[6]Getz (1993) defines tourist districts as areas with high concentrations of entertainment businesses and visitor-oriented attractions and services located in conjunction with central business districts.

[7]It should be noted more recent city zoning policies have adopted mixed use zoning; zoing allowing both commerical and residenial use. This has mainly been in the downtown area and popular commercial districts.

[8]The model developed in this paper is also closely aligned with the work of Fujita (1988).

[9]Since accommodations provided through Airbnb are typically homes and apartments in residential areas a tourist who chooses to stay in an Airbnb can be thought of as being redistributed into the residential district.

The increase in entertainment within the area makes the residential district a more attractive place for tourists to stay, which means more tourists will use Airbnb, further increasing local demand for entertainment. If the cycle of increasing tourism and demand for entertainment is left unchecked it can lead to what is referred to in the literature as tourism gentrification.[10]

While tourism gentrification is well established within the literature few works have looked to connect it with Airbnb.[11] Among the works that do draw this connection many are limited to case studies (Cócola-Grant, 2018; Ioannides et al., 2018; Sans and Domínguez, 2016). By analyzing data on multiple US cities this paper is able to extend this body of work by providing evidence of a systemic relationship between Airbnb and one part of tourism gentrification, the development of businesses.

The empirical analysis presented in this paper is motivated by two conjectures obtained from the theoretical model, which was developed by repurposing and extending Krugman's 1980 model of trade. The conjectures suggest as more tourists populate the residential area entertainment firms will enter and service firms will be crowded out due to rising fixed costs. Using data collected from a variety of sources including InsideAirbnb, the American Community Survey (ACS), and the US Census a reduced form model is estimated to determine the validity of the conjectures. Two primary results are able to be attained from the estimations. First, increasing Airbnb usage increases the number of entertainment firms. Second, Airbnb usage has does not have a statistically significant effect on the number of service firms.

The remainder of the paper is structured as follows. Section 2 describes the theoretical model and conjectures. Section 3 details the estimation strategy that will be implemented and describes the data. Section 4 presents the empirical results. Finally, Section 5 concludes.

2 Theoretical Model

Suppose a city has two product sectors, an entertainment (E) sector and a service (S) sector. Both sectors are assumed to have a large variety of differentiate goods, so many that the product space for each sector can be represented a continuum of products. Examples of entertainment firms/products include bars, restaurants, and other attractions, while service firms/products can be thought of as barbers, hardware stores, laundromats, etc. The city is assumed to be made up of two districts, a downtown district, district 1, and residential district, district 2. Firms are able to occupy either district, and the distribution of firms across the two districts will be determined endogenously.

In addition there are two types of consumers, tourists (T) and non-tourists (N). Tourists can be interpreted as consumers who do not reside in the city and are visiting for a short period of time. Since tourists are not spending an extend time within the city they will have little need for the products produced by the service sector, and will therefore predominantly consume good from the entertainment sector. Conversely, non-tourists are consumers who are residents of the

[10]Gotham (2005) defines "tourism gentrification" as the "...transformation of a middle-class neighbourhood into a relatively affluent and exclusive enclave marked by a proliferation of corporate entertainment and tourism venues."

[11]See *Tourism and Gentrification in Contemporary Metropolises* for more details about tourism gentrification and a discussion of the literature.

city or staying within the city for an extended period of time. As a result non-tourists will have a stronger preference service sector products, though they will still consume some goods produced by the entertainment sector. The distribution of consumers across districts will be determined exogenously.[12] Tourists and non-tourists in district j have Cobb-Douglas tastes over the two sectors given by the respective utility functions

$$U_j^T = C_{Ej}^{\mu} C_{Sj}^{1-\mu} \qquad U_j^N = C_{Ej}^{\lambda} C_{Sj}^{1-\lambda},$$

where μ (λ) represents the expenditure share on entertainment goods by tourists (non-tourists), $\mu > \lambda$, and C_{ij} represents a composite index of the consumption of sector $i \in \{E, S\}$ available to consumers in district j. The quantity index, C_{ij}, is a sub-utility function defined over the continuum of varieties available to consumers in district j. Let $c_{ijk}(\omega)$ denote the consumption of each available variety of sector i good in district j from district k, and n_{ik} is the range or "number" of varieties available in sector i in district k. Assume C_{ij} is defined by a CES function

$$C_{ij} = \left[\sum_{k=1}^{2} \int_0^{n_{ik}} c_{ijk}(\omega)^{\rho} d\omega \right]^{\frac{1}{\rho}}, \quad 0 < \rho < 1,$$

where $k \in 1, 2$ represent the two districts in the economy, the downtown and the residential districts.

Consumers are able to purchase products in either location; however, if a consumer from district j consumes a product from district k she will incur a transportation cost. The cost of transportation from district j to district k will be represented as a markup over the price of the good in district k. Therefore the price to a consumer in district j of consuming good i from district k will be the mill or f.o.b. price in district k times the transportation cost, $p_{ijk} = p_{ik}\tau_{jk}$ where $\tau_{jk} \geq 1$.[13] When $j = k$ $\tau_{jk} = 1$.

2.1 Consumer Problem

Consumers receive an exogenous income Y. Given Y, $p_{Ek}(\omega)$ for each entertainment firm, and $p_{Sk}(\omega)$ for each service firm the consumer's problem is to maximize utility subject to the budget constraint

$$Y = \sum_{k=1}^{2} \int_0^{n_{Ek}} p_{Ek}(\omega)\tau_{jk} c_{Ejk}(\omega) d\omega + \int_0^{n_{Sk}} p_{Sk}\tau_{jk}(\omega) c_{Sjk}(\omega) d\omega.$$

The consumer's problem can be solved in two steps. First, given the value of the composite for good i, C_{ij}, each $c_{ijk}(\omega)$ needs to be chosen so as to minimize the cost of attaining C_{ij}. Therefore,

[12] The exogeneity of the tourist distribution requires some additional assumptions. First, all tourists redistirbuted into the residential district are able to be accommidated. Second, non-tourists do not earn income from hosting a tourist. In reality, the movement of tourists is an endogenous process governed by the supply of and demand for Airbnb. However allowing the tourist distribution to be endogenous would necessate the addition of a market for Airbnb, which would substaintially complicate the model, and is left for future research.

[13] The transportation cost used in this model is similar to Samuelson's "iceberg" transporation costs. Consumers face a higher prices as a result of having to travel to consume the product. However, unlike in the traditional application of iceberg costs no physical product "melts away" or is lost in transport. Therefore, firms only have to produce what is actaully demanded. The "transportation" cost is intentded to capture the additional cost incurred from the inconvenience or time it take to travel to a different district.

consumers will solve the following minimization problem

$$\min \sum_{k=1}^{2} \int_{0}^{n_{ik}} p_{ik}(\omega)\tau_{jk}c_{ijk}(\omega)d\omega \quad \text{s.t.} \quad C_{ij} = \left[\sum_{k=1}^{2} \int_{0}^{n_{ik}} c_{ijk}(\omega)^{\rho}d\omega\right]^{\frac{1}{\rho}}$$

For simplicity assume all firms producing good i in district k sell their good for the same price, p_{ik}. Then the consumer's problem can be written as

$$\min \ \sum_{k=1}^{2} n_{ik}p_{ik}\tau_{jk}c_{ijk} \quad \text{s.t.} \quad C_{ij} = \left[\sum_{k=1}^{2} n_{ik}c_{ijk}^{\rho}\right]^{\frac{1}{\rho}}$$

First-order conditions to the expenditure minimization problem gives

$$\frac{c_{ijk}^{\rho-1}}{c_{ijl}^{\rho-1}} = \frac{p_{ik}\tau_{jk}}{p_{il}\tau_{jl}}.$$

Plugging this condition into the budget constraint and solving for c_{ijl} yields the compensated demand function for good i produced in district l and consumed in district j,

$$c_{ijl} = \frac{(p_{il}\tau_{jl})^{\frac{1}{\rho-1}}}{\left[\sum_{k=1}^{2} n_{ik}\left(p_{ik}\tau_{jk}\right)^{\frac{\rho}{\rho-1}}\right]^{\frac{1}{\rho}}}C_{ij}.$$

We can also derive an expression for the minimum cost of attaining C_{ij}. Expenditure on a single variety of sector i is $p_{il}\tau_{jl}c_{ijl}$, so using the above equation, summing over all varieties and summing over districts l gives

$$\sum_{l=1}^{2} n_{il}p_{il}\tau_{jl}c_{ijl} = \left[\sum_{k=1}^{2} n_{ik}\left(p_{ik}\tau_{jk}\right)^{\frac{\rho}{\rho-1}}\right]^{\frac{\rho-1}{\rho}} C_{ij}$$

The term multiplying C_{ij} on the right-hand side of the expression can be defined as the price index so that the price index times the quantity composite is equal to expenditure. Denote the price index for sector i in district j as $\mathbb{P}_{ij}$ gives

$$\mathbb{P}_{ij} = \left[\sum_{k=1}^{2} n_{ik}\left(p_{ik}\tau_{jk}\right)^{\frac{\rho}{\rho-1}}\right]^{\frac{\rho-1}{\rho}} = \left[\sum_{k=1}^{2} n_{ik}\left(p_{ik}\tau_{jk}\right)^{1-\sigma}\right]^{\frac{1}{1-\sigma}}$$

where $\rho \equiv \frac{\sigma-1}{\sigma}$. The price index, $\mathbb{P}_{ij}$, measures the minimum cost of purchasing a unit of the composite index C_{ij} of sector i. Demand for c_{ijk} can now be written more compactly as

$$c_{ijk} = \left(\frac{p_{ik}\tau_{jk}}{\mathbb{P}_{ij}}\right)^{\frac{1}{\rho-1}} C_{ij} = \left(\frac{p_{ik}\tau_{jk}}{\mathbb{P}_{ij}}\right)^{-\sigma} C_{ij} \tag{1}$$

The upper-level step of the consumer's problem is to divide total income between entertainment and service goods in aggregate, that is, to choose C_{Ej} and C_{Sj} so as to

$$\max \ U = C_{Ej}^{\mu}C_{Sj}^{1-\mu} \quad \text{s.t.} \quad Y = \mathbb{P}_{Ej}C_{Ej} + \mathbb{P}_{Sj}C_{Sj}.$$

Since tourists and non-tourists have comparable problems the analysis conducted for the remainder of this subsection will focus on tourists, but a similar strategy can be applied to the non-tourist's problem. The results of the first-order conditions from the problem above are $C_{Ej} = \mu \frac{Y}{\mathbb{P}_{Ej}}$ and $C_{Ej} = (1-\mu)\frac{Y}{\mathbb{P}_{Sj}}$. Plugging these solutions into equation (1) gives the following uncompensated demand functions for products produced in district k and consumed in district j

$$c_{Ejk} = \mu Y \frac{(p_{Ek}\tau_{jk})^{-\sigma}}{\mathbb{P}_{Ej}^{-(\sigma-1)}} \qquad c_{Sjk} = (1-\mu) Y \frac{(p_{Sk}\tau_{jk})^{-\sigma}}{\mathbb{P}_{Sj}^{-(\sigma-1)}}.$$

Summing across all locations in which the product is sold, total sales to tourists for a single location variety k, denoted q_{ik}^T, is given by

$$q_{Ek}^T = \mu Y \sum_{j=1}^{2} s_j^T \frac{(p_{Ek}\tau_{jk})^{-\sigma}}{\mathbb{P}_{Ej}^{-(\sigma-1)}} \qquad q_{Sk}^T = (1-\mu) Y \sum_{j=1}^{2} s_j^T \frac{(p_{Sk}\tau_{jk})^{-\sigma}}{\mathbb{P}_{Sj}^{-(\sigma-1)}}, \tag{2}$$

where s_j^T represents the share of the total population that is tourists in district j. The indirect utility function is able to be obtained by plugging the solutions from Eq (2) into the utility function,

$$U_j^T = \mu^{\mu}(1-\mu)^{1-\mu} Y \mathbb{P}_{Ej}^{-\mu} \mathbb{P}_{Sj}^{-(1-\mu)}.$$

The term $\mathbb{P}_{Ej}^{-\mu}\mathbb{P}_{Sj}^{-(1-\mu)}$ can be interpreted as the cost-of-living index for district j.

2.2 Firm Problem

Next, consider a particular firm producing a specific variety of sector i good in district j. Profits for this firm are given by

$$\pi_{ij} = p_{ij}Q_{ij} - cQ_{ij} - F,$$

where Q_{ij} is the total demand by tourists and non-tourists defined by Eq (2), $Q_{ij} = q_{ij}^T + q_{ij}^N$. Each firm is assumed to choose its price taking the price index, $\mathbb{P}_{ij}$, as given. Entertainment and service firms have symmetric problems so this analysis will focus on entertainment firms for the remainder of this subsection. Profit maximization implies

$$p_{Ej}^* = \left(\frac{\sigma}{\sigma-1}\right) c$$

for all varieties produced in location j.[14] Firms are allowed to freely enter and exit the market in response to profits or losses. Given the pricing rule derived above, the profits of a firm in district j are

$$\pi_{ij} = \frac{cQ_{ij}}{\sigma-1} - F.$$

[14] Derivation of the profit maximizing price can be found in the Appendix.

Imposing the zero-profit condition implies the equilibrium output of any active firm will be

$$Q_{ij}^{*} = \frac{F(\sigma-1)}{c}.$$

The equations for p_{ij}^{*} and Q_{ij}^{*} reveal the size of the market will not affect the markup of price over marginal cost nor the quantity at which individual goods are produced. Therefore, all market effects will work through changes in the number of varieties that are available. This result is an artifact of assuming demand has constant-elasticity of substitution along with assuming firms behave non-strategically, i.e. take the price index as given.[15]

To determine the optimal number of firms substitute the zero-profit quantity, Q_{ij}^{*}, the optimal price, p_{ij}^{*}, and the equation for the price index into the equation for Q_{ij} to get

$$F\sigma = \frac{\mu Y s_1^T + \lambda Y s_1^N}{n_{E1} + n_{E2}\tau^{1-\sigma}} + \frac{\left(\mu Y s_2^T + \lambda Y s_2^N\right)\tau^{-\sigma}}{n_{E1}\tau^{1-\sigma} + n_{E2}}$$

for district 1 and

$$F\sigma = \frac{\left(\mu Y s_1^T + \lambda Y s_1^N\right)\tau^{-\sigma}}{n_{E1} + n_{E2}\tau^{1-\sigma}} + \frac{\mu Y s_2^T + \lambda Y s_2^N}{n_{E1}\tau^{1-\sigma} + n_{E2}}$$

for district 2. Setting these two equation equal to each other and solving yields

$$n_{E1}^{*} = \frac{\mu Y s_1^T + \lambda Y s_1^N}{F\sigma\left(1 + \frac{(\psi\tau^{1-\sigma}-1)}{(\tau^{1-\sigma}-\psi)}\tau^{1-\sigma}\right)} + \frac{\left(\mu Y s_2^T + \lambda Y s_2^N\right)\tau^{-\sigma}}{F\sigma\left(\tau^{1-\sigma} + \frac{(\psi\tau^{1-\sigma}-1)}{(\tau^{1-\sigma}-\psi)}\right)} \tag{3}$$

$$n_{E2}^{*} = \frac{\mu Y s_1^T + \lambda Y s_1^N}{F\sigma\left(\frac{(\tau^{1-\sigma}-\psi)}{(\psi\tau^{1-\sigma}-1)} + \tau^{1-\sigma}\right)} + \frac{\left(\mu Y s_2^T + \lambda Y s_2^N\right)\tau^{-\sigma}}{F\sigma\left(\frac{(\tau^{1-\sigma}-\psi)}{(\psi\tau^{1-\sigma}-1)}\tau^{1-\sigma} + 1\right)}, \tag{4}$$

where $\psi = \frac{\mu Y s_1^T + \lambda Y s_1^N}{\mu Y s_2^T + \lambda Y s_2^N}$. A similar solution method can be used to find the optimal number of service firms for district 1, n_{S1}^{*}, and district 2, n_{S2}^{*}. To ensure the number of firms in all locations is non-negative, $n_{i1}^{*} \geq 0$ and $n_{i2}^{*} \geq 0$,

$$\tau^{1-\sigma} < \psi < \tau^{\sigma-1}$$

and

$$\tau^{1-\sigma} < \phi < \tau^{\sigma-1}$$

where $\phi = \frac{(1-\mu)Y s_1^T + (1-\lambda)Y s_1^N}{(1-\mu)Y s_2^T + (1-\lambda)Y s_2^N}$.

2.3 Numerical Analysis

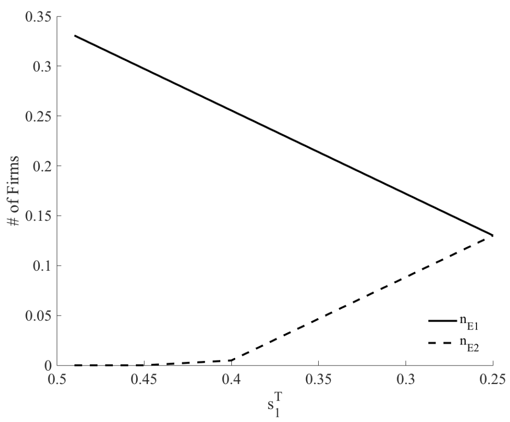
The purpose of this section is to better understand the relationship between the number of firms and the distribution of tourists by conducting a numerical analysis. Figure 1 plots equations (3) and (4) and the corresponding equations for the number of service firms as a function of the

[15] Relaxing either of these assumptions would allow the market size to have procompetitive effects. As more firms enter the market the price-cost margin would decrease implying firms would need to operate at a higher quantity in order to break even (Fujita et al., 2001). This extension is left for future research.

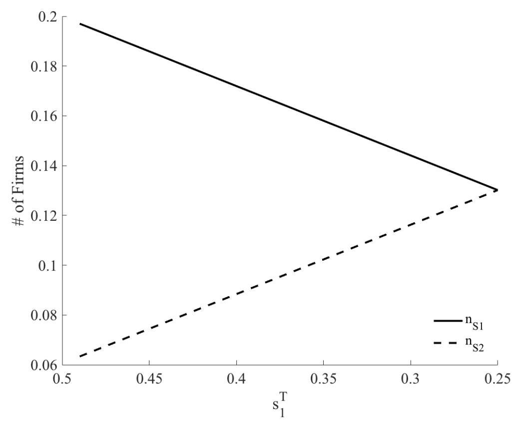
proportion of tourists in district one (downtown), s_1^T.[16] Assume there is a unit mass of consumers equally split between tourists and non-tourists, $s^T = 0.5$ and $s^N = 0.5$. Before Airbnb is introduced all the tourists reside downtown so $s_1^T = 0.5$ and $s_2^T = 0$. Figure 1a shows at this initial point all entertainment firms choose to locate downtown. As tourists start utilizing Airbnb s_1^T will decrease, s_2^T increases, and the number of entertainment firms downtown decreases while the number of entertainment firms in the residential district increases. A result inline with the Home Market Effect.

Conjecture 1. *Increasing the proportion of tourists within a district will lead to an increase in the number of entertainment sector firms.*

Figure 1



(a) Entertainment Sector



(b) Serivce Sector

Figure 1b shows the effect redistributing tourists has on the number of service firms. As the fraction of tourists downtown decreases the number of service firms also decreases. Conversely the number of service firms in the residential district increases in response to an increase in the local demand. Although tourists do not primarily consumer service sector products they do consume some, and since the number of residence in the district remains constant aggregate local demand will increase. This result is inline with the Home Market Effect, but does not agree with the observations about tourism gentrification made by Gotham (2005) and Cócola-Grant (2018). For the model to exhibit tourism gentrification the number of service firms in the residential district would need to decline as the presence of tourists, and thus entertainment firms, increased.

One explanation the result for service firms is not in line with the tourism gentrification process is the model fails to incorporate the effects of congestion, land scarcity, and rising rents. By holding fixed costs constant the model assumes a district is able to accommodate an infinite number of firms at no additional cost, which is not realistic. Land rents rise with demand (Fujita, 1988; Sivitanidou and Wheaton, 1992; Anas and Xu, 1999). Therefore, as the number of firms in a district increases, and with it demand for retail space, the fixed cost of producing in the region will also increase.

[16] Other parameters are assumed to be $s_1^N = s_2^N = 0.25$, $\mu = 0.75$, $\lambda = 0.25$, $t = 1.5$, $\sigma = 3.5$, $F = 1$, and $Y = 2$.

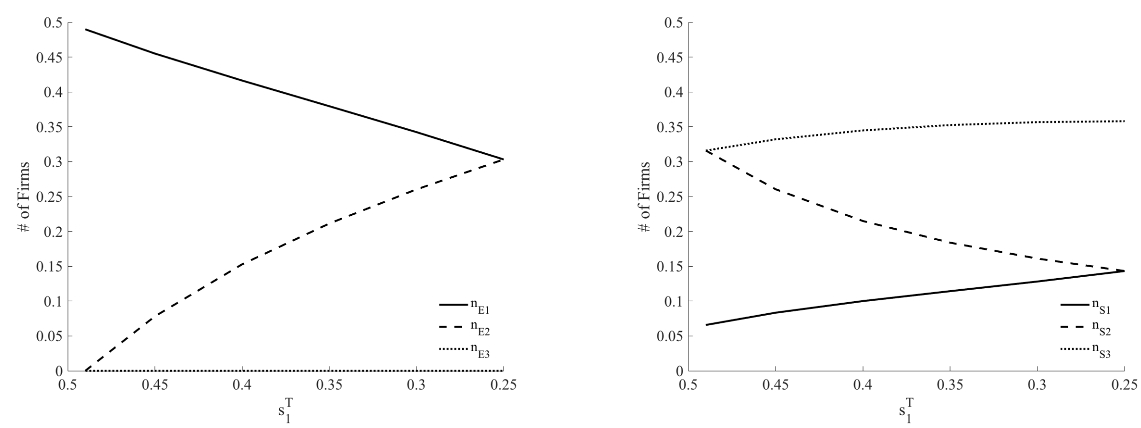
The next section extends the model to allow fixed costs to be a function of the number of firms and analyzes how the change impacts the results.

2.4 Endogenous Fixed Costs with 3-regions

Following the literature on urban economics, fixed costs for district j are assumed to be an increasing function of the number of firms within the district, $F_j(n_{Ej}, n_{Sj})$. Additionally a second residential district, district 3, is added to the model to represent the part of a city where Airbnb is not adopted. The fraction of tourists within this district is held constant at zero. By introducing a third district the model is able to also show what effects redistributing tourists has on areas not direction effected by the redistribution. For simplicity the function for fixed costs is assumed to be linear, $F_j = n_{Ej} + n_{Sj}$.

Figure 2 plots the number of firms for each district as a function of the number of tourists in district one (downtown), s_1^T.[17] The mass of consumers is again assumed to be one and equally split between tourists and non-tourists. The plot of entertainment firms in Figure 2a tells a similar story to what was seen in Figure 1a. When tourists reside in only the downtown district all entertainment firms choose to locate downtown. Then as tourists are redistributed to one of the residential districts, in this case district two, the number of entertainment firms in that residential district increases, while the number of entertainment firms downtown decreases. Additionally, the number of entertainment firms in the other residential district, district three, remains at zero since the local demand for entertainment firms is relatively small. The results in Figure 2a are in line with the Home Market Effect. As the local demand for entertainment firms increases (decrease) the number of entertainment firms increases (decreases).

Figure 2



(a) Entertainment Firms (b) Serivce Firms

In contrast, the effect of redistributing tourists on the number of service firms within a district now runs counter to the Home Market Effect. As tourists are redistributed into district two the number of service firms in the district declines even though local aggregate demand for service

[17]The system of equations used to solve for the number of firms can be found in the Appendix.

firms is increasing. This result occurs because fixed costs are increasing, due to an increase in the number of entertainment firms, at a faster pace than revenue. Therefore, profit maximizing service firms will choose to exit the more expensive district two and relocate in one of the other, cheaper districts. The result of Airbnb redistributing tourists into residential districts is the growth of firms in the entertainment sector causing firms in the service sector to be crowed out, also known as tourism gentrification.

Conjecture 2. *Increasing the proportion of tourists within a district will lead to a decrease in the number of service sector firms when the increase in fixed costs outweighs the increase in revenue.*

3 Empirical model and estimation

The theoretical model presented in the previous section yields two main conjectures about how redistributing tourists affects within a local community. First, as more tourists enter the residential district the number of entertainment sector firms will increase due to an increase in local aggregate demand. Second, service sector firms in the residential area are crowed out as a result of fixed costs increasing faster than revenue. The empirical analysis presented in this section analyzes the validity of these conjectures for 248 Public Use Microdata Areas (PUMA) across the US from 2005 to 2016.[18]

3.1 Basic Empirical Model

To investigate the questions proposed by the theoretical model, the following reduced-form specification for the number of firms within a PUMA is developed:

$$n_{ijt} = \mu_{ij} + \beta_{1i} s_{jt-1} + \beta_{2i} X_{jt-1} + \epsilon_{ijt}, \tag{5}$$

where i denotes the business sector, $i \in \{E, S\}$, j denotes the PUMA, and t denotes time in years. The variable n_{ijt} denotes the number of businesses in sector i within PUMA j at time t. s_{jt-1} denotes the measure of tourism within a PUMA from the pervious year. Tourism data is routinely available at the city level; however, tourism data at a more disaggregated level is not as common. Therefore instead of using explicit measures of tourism, data on Airbnb usage is employed as a proxy. The variable X_{jt-1} denotes a vector of PUMA specific time-varying controls including the total population, median income, and a count of hotels from the pervious year. Finally, μ_{ij} denotes PUMA fixed effects for sector i, the β's are the parameters to be estimated, and ϵ_{ijt} is the idiosyncratic error.

A firm's entry decision is inherently a slow moving process. It takes time for a firm to find a location and acquire the necessary inputs. Consequently, a firm deciding the market conditions are right for entry at time $t-1$ will likely not be observed entering the market until time t. For this

[18] Data is aggregated at the PUMA level since this is the smallest geographic unit made publicly available by the Census. A PUMA is defined by the Census as a statistical geographic area nested within states, containing at least 100,000 people, are built on census tracts and counties, and are geographically contiguous.
Source: `https://www.census.gov/geo/reference/puma.html`

reason lag, rather than contemporaneous, measures of the independent variables are included in the reduced form model.

3.2 Estimation strategy

Standard fixed effects methodology can be implemented to estimate Eq (5) , which allows the model to control for unobserved heterogeneity across PUMAs. Additionally using fixed effects allows the unobserved heterogeneity to be correlated with the regressors. For example, median income and total population are likely correlated with unobserved attributes of the PUMA, such as the quality of the area. Not accounting for this relationship would bias the estimates.[19] By including fixed effects the model accounts for the endogeneity caused by the unobserved heterogeneity; however it does not control for other sources of endogeneity.

Where tourists choose to locate will be, at least partially, determined by the types of firms within the area. Furthermore, the types of firms within the area depends on where tourists choose to locate. The cyclical nature of this problem leaves the model open to endogeneity bias that is not removed by the inclusion of fixed effects.

Since tourism within a PUMA depends on the types of firms within the PUMA, and the types of firms within the PUMA depends on tourism there is potential for endogeneity that is not removed with the inclusion of fixed effects. Let s_{jt} be represented by the following reduced form equation

$$s_{jt} = \nu_j + \beta_3 n_{Ejt} + \beta_4 n_{Sjt} + \beta_5 Z_{jt} + \gamma_{jt}, \tag{6}$$

where n_{Ejt} is the number of entertainment sector firms and n_{Sjt} is the number of service sector firms within PUMA j at time t, Z_{jt} is a vector of PUMA level observable characteristics, ν_j denotes PUMA fixed effects, and γ_{jt} represents the error term. Unlike Eq (5), the measure of Airbnb usage, s_{jt}, depends on contemporaneous regressors. A survey conducted by Gitelson and Crompton (1983) shows over 70% of individuals plan a vacation less than three months in advance which suggests current characteristics of an area will influence their location decision. Therefore, the reduced form model of Airbnb usage at time t depends on PUMA level characteristics and sector firm counts at time t.

While this paper is not interested in estimating Eq (6) it can be used to provide conditions for the exogeneity of s_{jt}, $\mathbb{E}[\epsilon_{ijt} s_{jt-1}] = 0$. Substituting Eq (6) into the equation of strict exogeneity of s_{jt-1} yields

$$\mathbb{E}[\epsilon_{ijt}(\nu_j + \beta_3 n_{ijt-1} + \beta_4 Z_{jt-1} + \gamma_{jt-1})] = 0,$$

which will hold as long as (i) the errors in the two equations are independent, $\mathbb{E}[\epsilon_{ijt}\gamma_{jt-1})] = 0$, (ii) the error ϵ_{ijt} is independent of PUMA fixed effects, $\mathbb{E}[\epsilon_{ijt}\nu_j] = 0$, (iii) the error ϵ_{ijt} is independent of the explanatory variables in Z_{jt-1}, $\mathbb{E}[\epsilon_{ijt} Z_{jt-1}] = 0$, and (iv) the errors do not exhibit autocorrelation, $\mathbb{E}[\epsilon_{ijt}\epsilon_{ijt-1}] = 0$.[20] Results of both the Arellano and Bond (1991) and Wooldridge (2010) test for serial correlation reject the null hypothesis of no autocorrelation. The strict exogeneity assumption on s_{jt-1} fails.

[19]Results of the Hausman test suggest a fixed effects specification is preffered to random effects which supports the claim observed regressors are correlated with unobserved heterogeneity.

[20]The final condition was derove by substituting a lagged version of Eq (5) into $\mathbb{E}[\epsilon_{ijt} n_{ijt-1}] = 0$.

A common method for addressing endogeneity is to estimate the model using an instrumental variable (IV) approach. The percentage of households with internet will be used as an instrument for Airbnb usage. To host an accommodation an individual needs to have access to Airbnb's website. If an individual has access to internet within the home it more convenient for the individual to post a listing, and as a result it is more likely the individual will choose listing their property. Furthermore, as the number of listing within an area increases so too will Airbnb usage.[21] Therefore household internet penetration is positively correlated with Airbnb usage.

Conversely, a firm's entry decision does not depend on household internet penetration. The types of firms being studied in this paper are those that provide a good or service that must be consumed in person. An individual cannot go to a restaurant or bar online. Additionally, traditional brick and mortar retailers as well as service providers like barbers and lawyers still require individuals to visit the physical location in order to consumer their product. Thus the entry decision for these firms will depend on the fixed cost and demographics of the PUMA, which are independent of internet accessibility.

It should be noted there does exist a subset of firms whose entry decision is likely to be depended on internet accessibility. Firms providing a good or service that can be consumed online. Think online retailers and service providers like Amazon and Goggle. However since their products can be consumed or purchased online these firms do not need to be in the immediate vicinity of the consumers. Accordingly, any changes in local demand, like one that would occur from an increase in tourism, will have no effect on the location decision of online retailers. For this reason these firms are not included in either the entertainment or service sector count of firms.[22]

A potential criticism of using internet penetration as an instrument for Airbnb usage is that it is a weak instrument, and therefore any estimates will suffer from weak instrument bias. Though the Kleibergen-Paap Wald rk F statistic rejects the weak instrument hypothesis, the correlation between the Airbnb usage and internet penetration is low, $R = 0.15$. Moreover, data on internet penetration is only available from 2013 to 2016. The shorted time series could further bias estimates. Ideally the robustness of the results would be checked by using an alternative instrument, but a strong external instrument is difficult to find. A natural next step is to look within the dataset to generate instruments via lags of the endogenous regressors; however, by demeaning the data lagged values of Airbnb usage become embedded in the transformed error and therefore are invalid instruments.

Alternatively first differences can be used to transform the data which will remove the fixed effects and avoid making lagged values of Airbnb usage endogenous.[23] First differencing does however result in the differenced error, $\Delta\epsilon_{ijt}$, no longer being i.i.d. Consequently using 2SLS

[21] Though more listings within an area does not necessitate higher levels of Airbnb activity, based on the Airbnb data collected the correlation coefficient between the number of listings and number of reviews, a proxy for usage, within a PUMA is $R = 0.8710$.

[22] Online retailers have a unique NAICS code, 454110, which allows them to be separetly identified from brick and morter retailers.

[23] The two most common transformations used when instrumenting with lags in the presence of fixed effects are first differencing and forward orthogonal projections (Arellano and Bover, 1995). Arellano and Bover show with balanced panels any two transformations of full row rank will yeild numerically identical estimators, holding the instrument set fixed. The panel used in this paper is unbalanaced, but only for one PUMA, PUMA 2400 in New Orleans, that is missing two years of data. Results are robust to the use of the forward orthogonal transformation.

results in distorted estimates. Following the work of Arellano and Bond (1991) and Arellano and Bover (1995)/Blundell and Bond (1998), more efficient, better behaving estimates can be achieved using system Generalized Method of Moments (GMM).

System GMM utilizes moment conditions based on lagged levels, $E(s_{jt-L}\Delta\epsilon_{ijt}) = 0$ for $t \geq 3$ and $L \geq 2$, as well as lagged first differences, $E(\Delta s_{jt-1}\epsilon_{ijt}) = 0$ for $t \geq 3$. Including both sets of moment conditions improves the estimates by taking advantage of more information, but also opens up the model to possibly overfitting endogenous variables. The Hansen test can be employed to diagnose any overfitting. Following the work of Roodman (2009), if the Hansen test has a p-val of at least 0.25 then it can be concluded the endogenous variables are not being overfit.[24]

3.3 Variable constuction and data

This section discusses the construction of the variables used to estimate Eq (5) as well as the data sources. Table 1 summarizes variable descriptions and provides summary statistics.

Table 1

Variable	Description	Units	Mean	SD	Observations
Dependent variables					
n_{Ejt}	Number of entertainment firms	10s	70.69	52.32	2976
n_{Sjt}	Number of service firms	10s	143.70	86.56	2976
$\%n_{Ejt}$	Pct of entertainment firms	% point	19.90	3.84	2976
$\%n_{Sjt}$	Pct of service firms	% point	41.62	7.42	2976
Explanatory variables					
# of Reviews (s_{jt})	Number of Airbnb reviews	10s	70.31	261.62	2976
Pop_{total}	Total population	10,000s	14.04	3.30	2974
Y_{50}	Median income	$1,000s	36.63	12.89	2974
# of Hotels	Number of hotels	10s	1.73	2.08	2976
Instruments					
% of HH with Internet	Pct of households with internet	% point	78.20	10.15	992

Note: No data is reported for PUMA 2400 in New Orleans in 2006 and 2007. Regressions were run as unbalanced and balanced (without PUMA 2400) panels. There was no significant effect on the results.

3.3.1 Dependent variable

The annual counts and percentages of entertainment and service sector firms within a PUMA were constructed using data collected from the Census County Business Patterns (CPB) database.

[24] If the model does suffer from overfitting, Hansen test p-val of less than 0.25, then the lag length will be restricted in order to reduce the number of moment conditions.

Data was collected from 2005 to 2016 for sixteen major US cities.[25] The annual datasets pulled from the CPB report the number of businesses within a zip code by six-digit NAICS code.

The data was able to be converted from zip code level to PUMA level using a crosswalk file obtained from the Missouri Census Data Center.[26] The crosswalk file reported the proportion of each zip code that was contained within the relevant PUMAs. For example, 6% of zip code 10003 (New York) falls within PUMA 3807, 24% falls within PUMA 3808, 44% falls within PUMA 3809, and 25% falls within PUMA 3810. These proportions are multiplied by the zip code business counts and then totaled by PUMA.

This methodology implicitly assumes businesses within an NAICS are uniformly distributed across a zip code, which is not the case. As this paper has already pointed out, businesses tend to cluster by sector as a result of zoning and agglomeration effects. Ideally, each business would be able to be assigned to a PUMA based on its geographic location, but this is not possible with publicly available data.[27]

After the data was converted to the PUMA level the businesses were sorted into the entertainment sector, service sector, or neither using the six-digit NAICS code. The general definition of an "entertainment sector" firm is a firm producing a good or service that is consumed by both tourists and residents. Examples of entertainment firms include bars, movie theaters, restaurants, and retail clothing stores. The general definition of a "service sector" firm is a firm producing a good or service consumed primarily by residents. Examples of service firms include auto mechanics, grocery stores, retail furniture stores, and tutoring services. Table 5 in the Appendix provides a detailed list of how every NAICS code was classified. After assigning each NAICS code to a sector the data was aggregated by year, PUMA, and sector to create an annual count of the number of firms in each sector by PUMA. Percentages were calculated by dividing the count for each sector by the total number of firms within the PUMA.

The classification for every NAICS code was determined by manually going through the codes and classifying them based on the previously stated definitions. Although this procedure is slightly ad hoc, the literature on tourism gentrification and urban economics provides no operational definitions for "entertainment" or "service" sector on which to base the classification. Most of the literature on tourism gentrification discusses these sectors as abstract concepts, not seeking to quantify the effect of tourism gentrification, and thus have not had a reason to provide an operational definition. Therefore the classification system defined in Table 5 should be thought of as a first attempt at providing the literature with a operational definition of the entertainment and service sectors for the purpose of analyzing tourism gentrification.

[25] The cities include in the dataset are: Asheville, NC; Austin, TX; Boston, MA; Chicago, IL; Denver, CO; Los Angeles, CA; Nashville, TN; New Orleans, LA; New York, NY; Oakland, CA; Portland, OR; San Diego, CA; San Francisco, CA; Santa Cruz, CA; Seattle, WA; and Washington DC. These cities were correspond to the cities available on InsideAirbnb, and were the only US citys for which data was made available.

[26] See `http://mcdc.missouri.edu`.

[27] Alternatively, the Department of Housing and Urban Development has crosswalk files available for zip codes to census tract, which can then be converted to PUMAs, that reports proportions based on business addresses. The proportion is the ratio of business addresses within the overlapping zip code-tract part to the total number of businesses within the zip code. Though the uniformly distributed assumption still applies, weighting the proportions by business addresses more accurately represents how businesses are distributed. However, these crosswalk files are only made available starting in 2010. Therefore, for the entire dataset to be included in the analysis the Missouri Census Data Center crosswalk files must be used.

3.3.2 Airbnb Usage

Data on Airbnb was collected from Inside Airbnb, an independent data collection project that compiles information about Airbnb listings for public use. The project scrapes Airbnb's website across various major US and international cities for host level data.[28] The raw data files include information about when the host joined Airbnb, geographic coordinates for the listing, the listing availability over the next year, when the listing received reviews, as well detailed information about the listing's amenities.

Since Inside Airbnb collects data by scraping Airbnb's website the data includes only those hosts who are active at the time of the scrape. Additionally Airbnb does provide a record of hosts that have since removed their listing from the website. Therefore, hosts who have removed their listing before the time of the scrape will not be included in the dataset. Without access to a database of historical data, the snapshot provided by Inside Airbnb is the best available option for measuring the intensity of Airbnb usage.

With the data provided a count of all listings available in an area can be created, but would overstate Airbnb usage. Just because a property is listed does not mean that it is being rented. Furthermore, Airbnb listings in and of themselves do not cause tourism gentrification. Rather, tourism gentrification is a result of an increase in tourism, which occurs when more tourists are brought into the area. Therefore only those listings that are rented should be included in the measure. A more accurate measure of Airbnb usage can be generated by exploiting data on listing reviews.

Inside Airbnb provides a record of all reviews a listing receives as well as the date the review was posted.[29] Utilizing this record, an annual count of the number of reviews a listing receives is generated. It should be noted that guests are not required to submit a review after renting. As a result a listing maybe more active then what is reflected by the number of reviews. In addition reviews can only be posted by an individual who has rented the listing. Therefore, the measure of Airbnb usage generated using review data should be thought of as a lower bound of Airbnb usage. It is also assumed the incentive to provide a review is consistent across time and PUMA since there is no explicit benefit to the renter for providing a review.

Using the geographic coordinates reported for each listing and boundary shape files obtained for the Census website each listing was able to be assigned to a PUMA. The count of reviews was then aggregated by PUMA year. The resulting variable should be thought of as a lower bound estimate of the number of times Airbnb was rented within a PUMA during the year. More aptly this variable can be interpreted as the minimum number of tourists brought into the PUMA by Airbnb since at least one tourist will accompany every rental. Of course, it is possible for more than one tourist to stay in an Airbnb, and in fact many of the listings allow two or more people to stay. However, the review data does not identify how many individuals stayed during the rental

[28]Though tourism gentrification can effect any city this paper chooses to focus on US cities for two reasons. First, demographic controls and business count data are easier to acquire for US cities. Second, many international cities do not have the same zoning laws as the US. Rather than having separate zoning for residential and commerical areas, international cities more commonly implement mixed use zoning. As a result it is more difficult to define distinct "residential" and "tourist" zones within a city.

[29]Guests have fourteen days to post a review after checking out. So even though the date does not correspond to the exact date of rental it is a close approximation.

period. Therefore to be most conservative and not over estimate the effect the minimum number of individuals that could have used Airbnb during the year is used as the measure of Airbnb usage.

3.3.3 Other controls

Controls for total population, median household income, and the number of hotels are also included in the model. Data on total population and median household income were collected from the American Community Survey (ACS) and were available at the PUMA level. Data on the number of hotels was collected from the CBP database. All businesses with the four-digit NAICS code 7211 (Traveler Accommodations) are included in the hotel count. The same methodology to convert business sector counts from zip code level to PUMA level was applied to the variable for hotel count.

The instrumental variable, percentage of households with internet, was also collected from the ACS. However, the ACS did not start collecting data on internet usage until 2013. Therefore, all IV estimates are calculated using only data from 2013 to 2016.[30]

4 Emperical Results

Table 2 presents the estimates of Eq (5) with the number of entertainment firms, n_{Ejt}, as the dependent variable. Specifications (1) and (2) present the OLS and FE estimates, respectively; however, both equations suffer from endogeneity bias. Specifically Airbnb usage is likely to be correlated with any unobserved heterogeneity, represented by the fixed effects. For example, consider a PUMA experiences a negative idiosyncratic shock to the number of firms at $t-1$ for a reason not modeled, so that the shock appears in the error term. Note the shock is assumed to affect firms in both the entertainment and service sectors.[31] All else equal, the fixed effects for the PUMA - the deviation of its average unexplained number of firms from the sample average - will be lower.

Furthermore, according to Eq (6) Airbnb usage at $t-1$ will also be affected by a shock to the number of firms. Estimation of Eq (6) reveals the coefficient on the number of entertainment sector firms is positive, $\beta_2 > 0$, the coefficient on the number of service sector firms is negative, $\beta_3 < 0$, and the magnitude of the coefficient on entertainment sector firms is larger than the magnitude of the coefficient on service sector firms, $|\beta_2| > |\beta_3|$. Regression results Eq (6) can be found in Table 6 in the Appendix. Based on the regression results, a negative shock to the number of firms should result in a decline in Airbnb usage. Both Airbnb usage at $t-1$ and the fixed effects are lower as a result of the shock. The positive correlation between lagged Airbnb usage and the error results in an inflated estimate of the OLS coefficient on Airbnb usage.

[30] The Consumer Population Survey provides data on internet usage prior to 2013, but not at a level of dissaggregation useful for this study.

[31] Shocks like a change in tax or regulation typically effect all businesses regradless of industry. Therefore the number of firms in both the entertainment and serivce sectors will be impacted.

Table 2

# of Entertainment Firms (n_{Ejt})	(1) OLS	(2) FE	(3) IV	(4) GMM
# of Reviews	0.0547***	0.0251***	0.0203***	0.0258*
(s_{jt-1})	(0.0056)	(0.0082)	(0.0046)	(0.0153)
Pop_{total}	2.5611***	0.4183	-0.3412	4.3003***
	(0.2081)	(0.5829)	(0.3636)	(1.2385)
Y_{50}	1.1723***	0.4519**	-0.2766*	-1.6373*
	(0.0568)	(0.2268)	(0.1513)	(0.9075)
# of Hotels	13.7920***	17.3385***	-0.7442	12.4061**
	(0.3582)	(1.9554)	(1.5270)	(4.9642)
PUMA FE	No	Yes	Yes	Yes
Year FE	Yes	Yes	No	Yes
F	258.0035	21.7182	11.5635	
AR(2)				0.5622
Hansen Test p-val				0.5789
Instruments			1	33
Years	2005-2016	2005-2016	2014-2016	2005-2016
N Obs	2726	2726	744	2726

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ Robust standard errors clustered at the PUMA level are reported in parentheses.

Including PUMA fixed effects in the estimation draws the unobserved heterogeneity out of the error term, and removes the endogeneity bias caused by the correlation between Airbnb usage and the fixed effects. In spite of that, Airbnb usage is still correlated with the error due to the presence of autocorrelation. Under the within group transformation the lag of Airbnb usage becomes $s^*_{jt-1} = s_{jt-1} - \frac{1}{T-1}(s_{j2} + ... + s_{jT})$ while the error become $\epsilon^*_{ijt} = \epsilon_{ijt} - \frac{1}{T-1}(\epsilon_{ij2} + ... + \epsilon_{ijt})$. Again assume a negative shock to the number of firms at time $t-1$. The shock causes the transformed error at time t, ϵ^*_{ijt}, to increase, and s_{jt-1} will decline for the same arguments as before. The error term and Airbnb usage are negative correlated. As a result estimation using FE will underestimate the true effect of Airbnb usage.

Specification (3) attempts to correct for the endogeneity in the FE model by utilizing the percentage of households with internet as an instrument for Airbnb usage. The coefficient on Airbnb usage is similar to the FE estimate, but given FE underestimates the true effect the IV estimate is lower than expected. However, this outcome is not unexpected for a couple of reasons. First the percentage of households with internet is weakly correlated with Airbnb usage, which can lead to weak instrument bias. Second, the estimation only utilizes data from 2013 to 2016, a period well after Airbnb was established and widely adopted. Failing to account for the variation in the number of businesses during years prior to and shortly after the start of Airbnb will bias the estimate of the effect Airbnb usage has on the number of entertainment firms.

The final specification is GMM with lags as instruments for the endogenous regressors. The p-value for the Hansen test is above the threshold of 0.25, and thus cannot reject the null hypothesis of joint validity of the instruments. Furthermore, the coefficient on Airbnb usage is significant at the 10% level. The results of the empirical analysis suggest as Airbnb usage within a PUMA increases more entertainment firms will choose to enter the area, which supports the conjecture made by the theoretical model.

Result 1. *Increasing Airbnb usage within a PUMA leads to an increase in the number of firms in the entertainment sector.*

Table 3 presents the estimates with the number of service firms, n_{Sjt}, as the dependent variable. The OLS estimate for the coefficient on Airbnb usage, presented in the second column of the table, is again an overestimate of the true effect due to the positive correlation between the error term and Airbnb usage.[32] Similarly, the FE estimate, presented in column three of the table, again underestimates the true effect due to the negative correlation between the transformed regressor and transformed error. Furthermore, the effect is insignificant even at the 10% level. The IV estimate presented in the fourth column of the table is significant at the 1% level, but is larger the OLS estimate which is expected to overestimate the true effect. Again, the incompatibility of the result could be occurring because of weak instrument bias and restricted time series.

Table 3

# of Service Firms (n_{Sjt})	(1) OLS	(2) FE	(3) IV	(4) GMM
# of Reviews	0.0365***	0.0076	0.0383***	0.0142
(s_{jt-1})	(0.0099)	(0.0104)	(0.0101)	(0.0102)
Pop_{total}	5.7987***	1.6292	-0.2382	9.3962***
	(0.3658)	(1.0689)	(0.7505)	(2.9535)
Y_{50}	2.4193***	0.6874**	-0.5258	-3.2946
	(0.0999)	(0.3116)	(0.3313)	(2.6584)
# of Hotels	21.2947***	31.1340***	-1.7606	16.0910*
	(0.6297)	(7.3553)	(2.9777)	(9.1277)
PUMA FE	No	Yes	Yes	Yes
Year FE	Yes	Yes	No	Yes
F	233.4746	31.2784	8.4620	
AR(2)				0.9902
Hansen Test				0.6407
Instruments			1	21
Years	2005-2016	2005-2016	2014-2016	2005-2016
N Obs	2726	2726	744	2726

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ Robust standard errors clustered at the PUMA level are reported in parentheses.

[32] Although the coefficient on n_{Sjt} is negative when Airbnb usage is regressed on the number of firms it is smaller than the coefficient on n_{Ejt}, so there is a positive net effect to Airbnb usage when there is a positive shock to the number of firms.

Specification (4) of the table presents the results of the GMM estimation. The Hansen test is found to be insignificant, so the instruments can be regarded as jointly valid. Furthermore, the estimate on Airbnb usage is similar to the OLS and FE estimates. While there is support for the validity of the estimation strategy the coefficient on Airbnb usage is insignificant even at the 10% level. Thus the empirical results are unable to conclude Airbnb usage has any effect on the number of service firms within a PUMA.

Result 2. *Airbnb usage does not have a statistically significant effect on the number of service sector firms within a PUMA.*

The absence of an effect by Airbnb usage on the number of service firms contradicts Conjecture 2, increasing the proportion of tourists within a district will lead to a decrease in the number of service sector firms when the increase in fixed costs outweighs the increase in revenue. However a null effect could occur if the net effect of a change in fixed costs and a change in revenue was zero, which would happen with slight modifications to the assumptions of the theoretical model. The model assumes the utility of tourists depends on the service firm, which is not necessarily the case. Tourists typically visit a city or place for a short period of time, and as a result will primarily consume the products and services produced by the entertainment sector. The goods and services provided by the service sector are typically consumed only by those individuals staying within an area for an extended period of time. Therefore in an extreme case the utility of tourists can be assumed to put no weight on the consumption of service sector products.

In this case tourists will not consume any products produced by the service sector. Then as Airbnb usage increases, tourists are redistributed into the residential area, the local aggregate demand for service sector firms will not change, and new service sector firms will have no incentive to enter the market. On the other hand, firms in the entertainment sector will still experience an increase in local aggregate demand, and thus have incentive to enter the residential area. If the resulting increase in the number of entertainment sector firms is small enough such that the fixed cost of production is effectively unchanged then service sector firms will have no incentive to exit the market. The number of firm in the service sector will not change. However, if the increase entertainment sector firms does lead to an increase in fixed costs firms in the service sector will have an incentive to exit the market. When fixed costs increase without an offsetting incentive to enter the market an increase in Airbnb usage would lead to a decline in the number of service firms.

Therefore for the number of service sector firms to remain unchanged when fixed costs increase there must also be an increase in local aggregate demand, which will occur if the utility of tourists places a positive weight on service sector products. As more tourists enter the residential area aggregate demand for service sector firms increases resulting in an increase in revenue for the firms already in the market. At the same time fixed costs are also increasing as a result of higher demand for retail space. The net of these two forces will determine whether there is an increase, decrease, or no effect on the number of service firms.

In the numerical analysis conducted earlier the effect of rising fixed costs was assumed to dominate the increase in demand resulting in a net outflow of service sector firms as more tourists enter the residential area. If instead the marginal effect firms have on fixed cost is reduced the resulting outflow of service firms will also diminish. Moreover, the marginal effect can be small enough such that the incentive to exit from increased fixed costs is exactly offset by the increase in aggregate demand. The result in this case is no observable change in the number of service sector firms in response to an increase in Airbnb usage.

As a robustness check Eq (5) can be estimated with entertainment/service sector firms as a percentage of the total number of firms within a PUMA as the dependent variable. Based on the results of Table 2 and 3 the coefficient on Airbnb usage should be positive when the percentage of entertainment firms is the dependent variable and negative when the percentage of service firms is the dependent variable. Results 1 and 2 state an increase in Airbnb usage leads to an increase in the number of entertainment sector firms and no change in the number of service sector firms. Since the number of entertainment sector firms is growing while the number of service sector firms is unchanged the proportion of firms within a PUMA in the entertainment sector should increase. Conversely, since the number of service sector firms is unchanged while the number of entertainment sector firms increases the proportion of service sector firms in a PUMA should decrease. Table 4 presents the regression results when percentage of entertainment sector firms and percentage of service sector firms are the dependent variable.[33]

[33] A linear probability model is assumed to be a good approximation since the dependent variables are not close to the bounds, $\%n_{Ejt}, \%n_{Sjt} \in [0, 1]$. The min and max of the percentage of entertainment sector firms , $\%n_{Ejt}$, are 8.61% and 33.19%, respectively, and the min and max of the percentage of service sector firms , $\%n_{Sjt}$, are 14.32% 64.39%, respectively.

Table 4

	% of Entertainment Firms ($\%n_{Ejt}$)		% of Service Firms ($\%n_{Sjt}$)	
	IV	GMM	IV	GMM
# of Reviews	0.0008	0.0003**	-0.0002	-0.0028**
(s_{jt-1})	(0.0006)	(0.0001)	(0.0009)	(0.0013)
Pop_{total}	-0.0597	0.0129	0.0871	0.1430
	(0.0535)	(0.0533)	(0.0668)	(0.0911)
Y_{50}	-0.0151	-0.0038	0.0075	0.0210
	(0.0190)	(0.0176)	(0.0262)	(0.0344)
# of Hotels	0.1392	-0.0224	0.1468	-1.0197***
	(0.1367)	(0.1285)	(0.1765)	(0.1747)
PUMA FE	Yes	Yes	Yes	Yes
Year FE	No	Yes	No	Yes
F	1.9072		1.2555	
AR(2)		0.8059		0.8186
Hansen Test		0.5013		0.3206
Instruments	1	29	1	78
Years	2014-2016	2005-2016	2014-2016	2005-2016
N Obs	744	2726	744	2726

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ Robust standard errors clustered at the PUMA level are reported in parentheses.

The second and third columns of Table 4 show the IV and GMM results with the percentage of entertainment sector firms is the dependent variable. The coefficient on the number of Airbnb reviews is positive, as expected. The fourth and fifth columns of the table show the results when the percentage of service firms is the dependent variable, and the coefficient on Airbnb usage is negative, again as expected. Furthermore, the GMM coefficient is significant at the 5% level for both the percentage of entertainment sector and service sector firms. The estimates presented in Table 4 provide further support for the validity of Results 1 and 2.

5 Conclusion

An overlooked consequence of the introduction and proliferation of Airbnb has been the redistribution of tourists into parts of cities that perviously have had little exposure to tourism. This paper examines, both theoretically and empirically, the effects redistributing tourists has had on the development of the entertainment and services sectors within city neighborhoods. To investigate this question theoretically a model of intra-city trade based on work by Krugman (1980) was developed. Two conjectures were able to be drawn from the model. When tourists, entertainment centric consumers, are concentrated in the downtown area, as they have historically been, firms in the entertainment sector choose to agglomerate downtown. Then as tourists are redistributed into the residential district, as they are with Airbnb, the number of entertainment sector firms in residential district increases. Furthermore, the number of service sector firms in the residential district will decrease if fixed costs (rents) increase, due to an increased demand for retail space, more than revenue increases.

An empirical analysis was then conducted to test the validity of the conjectures. For this analysis a novel data set was created by combining data from InsideAirbnb as well as a variety of Census sources. Data was collected from 2005-2016 at the PUMA level. The empirical results suggest Airbnb usage does have an effect the development of businesses within a PUMA, which led to two primary results. First, an increase in Airbnb usage results in an increase in the number of entertainment sector firms. Second, the data does not provide sufficient evidence to suggest Airbnb usage has a statistically significant effect on number of service sector firms. A result that can be explained theoretically by the increase in aggregate demand exactly offsetting the increase in fixed costs, when means there is no incentive for service sector firms to enter or exit the market. Thus there will be no observable effect on the number of firms in the service sector.

There are some limitations to the analysis and data. First, there is no operational definition of "entertainment" and "service" firms within the existing literature. The definition provided in this paper is a first pass at defining these two sectors, and a more rigorous definition should still be pursued. Second, Airbnb is not the only way tourists are able to reside outside of a city's downtown district. Cities have hotels outside of the central business district and downtown that will redistribute tourists into the residential area. The empirical model does attempt to account for this by including a count of the number of hotels, but this may not fully capture the effect alternative accommodations have on the business sector within a PUMA. Finally, tourists are not confined to the immediate area around their accommodation. Rather they are able to travel to different parts of a city, which means the economic impact of tourists may not correspond to where tourists stay. Though without detailed data on tourist flows within a city it is not possible to account for the effects tourists have on areas away from their accommodation. Additionally, research by Versichele et al. (2014) and Shoval et al. (2011) suggests tourists typically spend most of their time and money around the area of their accommodation.

Nevertheless, the results of this paper have implications that bridge the empirical literature on Airbnb and tourism gentrification. Prior to this paper the research on Airbnb has been largely limited to analyzing the effects Airbnb has had on hotel revenues and the housing market, but this paper suggests Airbnb may also impact the development of a district's entertainment and service sectors. Specifically, an increase in Airbnb usage leads to an increase in the presence of entertainment sector firms. A process referred to in the literature as tourism gentrification. Work by Cócola-Grant (2016; 2018) provides anecdotal evidence and motivation for the connection between the growth of Airbnb and tourism gentrification, but is limited to a case study of Barcelona. By focusing in on a singular aspect of tourism gentrification, the presence of entertainment sector firms, the results of this paper support and extend the work of Cócola-Grant by providing evidence of a systematic link between the increase in Airbnb usage and increase in entertainment sector firms. However, this paper is not able to definitively conclude Airbnb usage will lead to tourism gentrification.

There are other aspects of tourism gentrification not explored within this paper that are needed to conclude a causal link exists. Like gentrification in general, tourism gentrification can lead to residential displacement as a result of higher rents, as well as displacement resulting from long-term residential accommodations being converted to exclusive short-term tourist accommodations. Some states have begun addressing the latter issue by implementing policies restricting Airbnb hosts to listing only their primary residence or requiring hosts obtain a short-term rental license.[34] While these types of policies may help prevent the decline in availability of residential accommodations the potential issue of rising rents is still largely unaddressed. Additionally, congestion and the replacement of locally own firms with national chains also definitive characteristics of tourism gentrification. Though this paper is not able to conclude a causal link between Airbnb and tourism gentrification exits, it does demonstrate a connection between Airbnb usage and the presence of entertainment sector firms, which provides motivation for further research into the role Airbnb plays in the tourism gentrification process.

[34] Airbnb provides detailed informaiton about regulations in various cities. Additionally, Cohen (2018), Honan (2018), and Loudenback (2018) discuss various regualtions being imposed by states.

6 Appendix

Derive compensated demand

Using the FOC from the consumer's expenditure minimization problem the following equation for the marginal rate of substitution can be obtained

$$\frac{c_{ijk}^{\rho-1}}{c_{ijl}^{\rho-1}} = \frac{p_{ik}\tau_{jk}}{p_{il}\tau_{jl}}.$$

$$c_{ijk} = c_{ijl}\left(\frac{p_{il}\tau_{jl}}{p_{ik}\tau_{jk}}\right)^{\frac{1}{1-\rho}}$$

Plugging the above equation into the budget constraint yields

$$C_{ij} = \left[\sum_{k=1}^{2} n_{ik}\left(c_{ijl}\left(\frac{p_{il}\tau_{jl}}{p_{ik}\tau_{jk}}\right)^{\frac{1}{1-\rho}}\right)^{\rho}\right]^{\frac{1}{\rho}}$$

$$C_{ij} = c_{ijl}(p_{il}\tau_{jl})^{\frac{1}{1-\rho}}\left[\sum_{k=1}^{2} n_{ik}\left(\frac{1}{p_{ik}\tau_{jk}}\right)^{\frac{\rho}{1-\rho}} d\omega\right]^{\frac{1}{\rho}}$$

Solve for c_{ijl}

$$c_{ijl} = \frac{(p_{il}\tau_{jl})^{\frac{1}{\rho-1}}}{\left[\sum_{k=1}^{2} n_{ik}\left(p_{ik}\tau_{jk}\right)^{\frac{\rho}{\rho-1}}\right]^{\frac{1}{\rho}}} C_{ij}$$

Solve for profit maximizing price

Taking the derivative of the profit function with respect to p_{Ej} yields

$$\frac{\partial \pi_{Ej}}{\partial p_{Ej}} = Q_{Ej} + (p_{Ej} - c)\frac{\partial Q_{Ej}}{\partial p_{Ej}} = 0$$

Substitute in the definition of Q_{Ej}

$$q_{Ej}^{T} + q_{Ej}^{N} = (c - p_{Ej})\left(\frac{\partial q_{Ej}^{T}}{\partial p_{Ej}} + \frac{\partial q_{Ej}^{N}}{\partial p_{Ej}}\right)$$

Calculate the derivate of q_{Ej}^{T} and q_{Ej}^{N} with respect to p_{Ej}

$$\frac{\partial q_{Ej}^{T}}{\partial p_{Ej}} = \mu Y(-\sigma)p_{Ej}^{-1}\sum_{k=1}^{2} s_{k}^{T}\frac{(p_{Ej}\tau_{kj})^{-\sigma}}{\mathbb{P}_{Ej}^{-(\sigma-1)}}$$

$$\frac{\partial q_{Ej}^{N}}{\partial p_{Ej}} = \lambda Y(-\sigma)p_{Ej}^{-1}\sum_{k=1}^{2} s_{k}^{N}\frac{(p_{Ej}\tau_{kj})^{-\sigma}}{\mathbb{P}_{Ej}^{-(\sigma-1)}}$$

Plug in the equations for q^T_{Ej} and q^N_{Ej} as well as the equations for $\frac{\partial q^T_{Ej}}{\partial p_{Ej}}$ and $\frac{\partial q^N_{Ej}}{\partial p_{Ej}}$

$$\begin{aligned}\mu Y \sum_{k=1}^{2} s^T_k \frac{(p_{Ej}\tau_{kj})^{-\sigma}}{\mathbb{P}_{Ej}^{-(\sigma-1)}} + \lambda Y \sum_{k=1}^{2} s^N_k \frac{(p_{Ej}\tau_{kj})^{-\sigma}}{\mathbb{P}_{Ej}^{-(\sigma-1)}} \\ = (c - p_{Ej}) \left(\mu Y(-\sigma) p_{Ej}^{-1} \sum_{k=1}^{2} s^T_k \frac{(p_{Ej}\tau_{kj})^{-\sigma}}{\mathbb{P}_{Ej}^{-(\sigma-1)}} + \lambda Y(-\sigma) p_{Ej}^{-1} \sum_{k=1}^{2} s^N_k \frac{(p_{Ej}\tau_{kj})^{-\sigma}}{\mathbb{P}_{Ej}^{-(\sigma-1)}} \right)\end{aligned}$$

Solve the above equation for p_{Ej}.

$$\begin{aligned}p_{Ej} &= (p_{Ej}\sigma - c\sigma) \\ p_{Ej}(1-\sigma) &= -c\sigma \\ p^*_{Ej} &= \left(\frac{\sigma}{\sigma-1}\right) c\end{aligned}$$

Solve for optimal number of entertainment sector firms

Plug the equations for Q_{Ej}, q^T_{Ej}, and q^N_{Ej} into the entertainment sector market clearing condition.

$$\begin{aligned}Q_{Ej} &= q^T_{Ej} + q^N_{Ej} \\ \frac{F(\sigma-1)}{c} &= \mu Y \sum_{k=1}^{2} s^T_k \frac{(p_{Ej}\tau_{kj})^{-\sigma}}{\mathbb{P}_{Ek}^{-(\sigma-1)}} + \lambda Y \sum_{k=1}^{2} s^N_k \frac{(p_{Ej}\tau_{kj})^{-\sigma}}{\mathbb{P}_{Ek}^{-(\sigma-1)}}\end{aligned}$$

Expand the summations.

$$\begin{aligned}\frac{F(\sigma-1)}{c} &= \mu Y \left(s^T_1 \frac{(p_{Ej}\tau_{1j})^{-\sigma}}{\mathbb{P}_{E1}^{-(\sigma-1)}} + s^T_2 \frac{(p_{Ej}\tau_{2j})^{-\sigma}}{\mathbb{P}_{E2}^{-(\sigma-1)}} \right) + \lambda Y \left(s^N_1 \frac{(p_{Ej}\tau_{1j})^{-\sigma}}{\mathbb{P}_{E1}^{-(\sigma-1)}} + s^N_2 \frac{(p_{Ej}\tau_{2j})^{-\sigma}}{\mathbb{P}_{E2}^{-(\sigma-1)}} \right) \\ \frac{F(\sigma-1)}{c} &= (\mu Y s^T_1 + \lambda Y s^N_1) \frac{(p_{Ej}\tau_{1j})^{-\sigma}}{\mathbb{P}_{E1}^{-(\sigma-1)}} + (\mu Y s^T_2 + \lambda Y s^N_2) \frac{(p_{Ej}\tau_{2j})^{-\sigma}}{\mathbb{P}_{E2}^{-(\sigma-1)}} \\ \frac{F(\sigma-1)}{c} &= (\mu Y s^T_1 + \lambda Y s^N_1)(p_{Ej}\tau_{1j})^{-\sigma}\mathbb{P}_{E1}^{\sigma-1} + (\mu Y s^T_2 + \lambda Y s^N_2)(p_{Ej}\tau_{2j})^{-\sigma}\mathbb{P}_{E2}^{\sigma-1}\end{aligned}$$

Substitute in the equation for the optimal price, p^*_{Ej}.

$$\frac{F(\sigma-1)}{c} = \left(\frac{\sigma}{\sigma-1}c\right)^{-\sigma} (\mu Y s^T_1 + \lambda Y s^N_1)\tau_{1j}^{-\sigma}\mathbb{P}_{E1}^{\sigma-1} + (\mu Y s^T_2 + \lambda Y s^N_2)\tau_{2j}^{-\sigma}\mathbb{P}_{E2}^{\sigma-1}$$

Substitute in the definition of the price index, $\mathbb{P}_{Ej}$.

$$\begin{aligned}\frac{F(\sigma-1)}{c} = \left(\frac{\sigma}{\sigma-1}c\right)^{-\sigma} \Bigg[\left(\mu Y s^T_1 + \lambda Y s^N_1\right) \tau_{1j}^{-\sigma} \left(\left[n_{E1}(p_{E1}\tau_{11})^{1-\sigma} + n_{E2}(p_{E2}\tau_{12})^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \right)^{\sigma-1} \\ + \left(\mu Y s^T_2 + \lambda Y s^N_2\right) \tau_{2j}^{-\sigma} \left(\left[n_{E1}(p_{E1}\tau_{21})^{1-\sigma} + n_{E2}(p_{E2}\tau_{22})^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \right)^{\sigma-1} \Bigg]\end{aligned}$$

$$\frac{F(\sigma-1)}{c} = \left(\frac{\sigma}{\sigma-1}c\right)^{-\sigma}\Bigg[\left(\mu Y s_1^T + \lambda Y s_1^N\right)\tau_{1j}^{-\sigma}\left(n_{E1}(p_{E1})^{1-\sigma} + n_{E2}(p_{E2}\tau_{12})^{1-\sigma}\right)^{-1} + \left(\mu Y s_2^T + \lambda Y s_2^N\right)\tau_{2j}^{-\sigma}\left(n_{E1}(p_{E1}\tau_{21})^{1-\sigma} + n_{E2}(p_{E2})^{1-\sigma}\right)^{-1}\Bigg]$$

The profit maximizing price is independent of location, $p^*_{E1} = p^*_{E2} = p$.

$$\frac{F(\sigma-1)}{c} = \left(\frac{\sigma}{\sigma-1}c\right)^{-\sigma} p^{\sigma-1}\Bigg[\left(\mu Y s_1^T + \lambda Y s_1^N\right)\tau_{1j}^{-\sigma}\left(n_{E1} + n_{E2}\tau_{12}^{1-\sigma}\right)^{-1} + \left(\mu Y s_2^T + \lambda Y s_2^N\right)\tau_{2j}^{-\sigma}\left(n_{E1}\tau_{21}^{1-\sigma} + n_{E2}\right)^{-1}\Bigg]$$

$$\frac{F(\sigma-1)}{c} = \left(\frac{\sigma}{\sigma-1}c\right)^{-1}\left[\frac{\left(\mu Y s_1^T + \lambda Y s_1^N\right)\tau_{1j}^{-\sigma}}{n_{E1} + n_{E2}\tau_{12}^{1-\sigma}} + \frac{\left(\mu Y s_2^T + \lambda Y s_2^N\right)\tau_{2j}^{-\sigma}}{n_{E1}\tau_{21}^{1-\sigma} + n_{E2}}\right]$$

$$F\sigma = \frac{\left(\mu Y s_1^T + \lambda Y s_1^N\right)\tau_{1j}^{-\sigma}}{n_{E1} + n_{E2}\tau_{12}^{1-\sigma}} + \frac{\left(\mu Y s_2^T + \lambda Y s_2^N\right)\tau_{2j}^{-\sigma}}{n_{E1}\tau_{21}^{1-\sigma} + n_{E2}}$$

The equation for district 1 is

$$F\sigma = \frac{\mu Y s_1^T + \lambda Y s_1^N}{n_{E1} + n_{E2}\tau_{12}^{1-\sigma}} + \frac{\left(\mu Y s_2^T + \lambda Y s_2^N\right)\tau_{21}^{-\sigma}}{n_{E1}\tau_{21}^{1-\sigma} + n_{E2}},$$

and the equation for district 2 is

$$F\sigma = \frac{\left(\mu Y s_1^T + \lambda Y s_1^N\right)\tau_{12}^{-\sigma}}{n_{E1} + n_{E2}\tau_{12}^{1-\sigma}} + \frac{\mu Y s_2^T + \lambda Y s_2^N}{n_{E1}\tau_{21}^{1-\sigma} + n_{E2}}.$$

The transportation cost from district 1 to district 2 is equal to the transportation cost from district 2 to district 1, $\tau_{12} = \tau_{21} = \tau$.

$$F\sigma = \frac{\mu Y s_1^T + \lambda Y s_1^N}{n_{E1} + n_{E2}\tau^{1-\sigma}} + \frac{\left(\mu Y s_2^T + \lambda Y s_2^N\right)\tau^{-\sigma}}{n_{E1}\tau^{1-\sigma} + n_{E2}} \qquad F\sigma = \frac{\left(\mu Y s_1^T + \lambda Y s_1^N\right)\tau^{-\sigma}}{n_{E1} + n_{E2}\tau^{1-\sigma}} + \frac{\mu Y s_2^T + \lambda Y s_2^N}{n_{E1}\tau^{1-\sigma} + n_{E2}}$$

Set the two equations above equal to each other.

$$\frac{\mu Y s_1^T + \lambda Y s_1^N}{n_{E1} + n_{E2}\tau^{1-\sigma}} + \frac{\left(\mu Y s_2^T + \lambda Y s_2^N\right)\tau^{-\sigma}}{n_{E1}\tau^{1-\sigma} + n_{E2}} = \frac{\left(\mu Y s_1^T + \lambda Y s_1^N\right)\tau^{-\sigma}}{n_{E1} + n_{E2}\tau^{1-\sigma}} + \frac{\mu Y s_2^T + \lambda Y s_2^N}{n_{E1}\tau^{1-\sigma} + n_{E2}}$$

$$\mu Y s_1^T + \lambda Y s_1^N + \left(\mu Y s_2^T + \lambda Y s_2^N\right)\tau^{-\sigma}\frac{n_{E1} + n_{E2}\tau^{1-\sigma}}{n_{E1}\tau^{1-\sigma} + n_{E2}} = \left(\mu Y s_1^T + \lambda Y s_1^N\right)\tau^{-\sigma} + \mu Y s_2^T + \lambda Y s_2^N\frac{n_{E1} + n_{E2}\tau^{1-\sigma}}{n_{E1}\tau^{1-\sigma} + n_{E2}}$$

$$(\mu Y s_1^T + \lambda Y s_1^N)(1 - \tau^{-\sigma}) = (\mu Y s_2^T + \lambda Y s_2^N)(1 - \tau^{-\sigma})\frac{n_{E1} + n_{E2}\tau^{1-\sigma}}{n_{E1}\tau^{1-\sigma} + n_{E2}}$$

$$\frac{\mu Y s_1^T + \lambda Y s_1^N}{\mu Y s_2^T + \lambda Y s_2^N} = \frac{n_{E1} + n_{E2}\tau^{1-\sigma}}{n_{E1}\tau^{1-\sigma} + n_{E2}}$$

$$\psi = \frac{n_{E1} + n_{E2}\tau^{1-\sigma}}{n_{E1}\tau^{1-\sigma} + n_{E2}}$$

$$\psi(n_{E1}\tau^{1-\sigma} + n_{E2}) = n_{E1} + n_{E2}\tau^{1-\sigma}$$

$$n_{E1}(\psi\tau^{1-\sigma} - 1) = n_{E2}(\tau^{1-\sigma} - \psi)$$

$$n_{E1} = n_{E2}\frac{(\tau^{1-\sigma} - \psi)}{(\psi\tau^{1-\sigma} - 1)}$$

$$F\sigma = \frac{\mu Y s_1^T + \lambda Y s_1^N}{n_{E2}\frac{(\tau^{1-\sigma}-\psi)}{(\psi\tau^{1-\sigma}-1)} + n_{E2}\tau^{1-\sigma}} + \frac{\left(\mu Y s_2^T + \lambda Y s_2^N\right)\tau^{-\sigma}}{n_{E2}\frac{(\tau^{1-\sigma}-\psi)}{(\psi\tau^{1-\sigma}-1)}\tau^{1-\sigma} + n_{E2}}$$

$$n_{E2}^* = \frac{\mu Y s_1^T + \lambda Y s_1^N}{F\sigma\left(\frac{(\tau^{1-\sigma}-\psi)}{(\psi\tau^{1-\sigma}-1)} + \tau^{1-\sigma}\right)} + \frac{\left(\mu Y s_2^T + \lambda Y s_2^N\right)\tau^{-\sigma}}{F\sigma\left(\frac{(\tau^{1-\sigma}-\psi)}{(\psi\tau^{1-\sigma}-1)}\tau^{1-\sigma} + 1\right)}$$

$$F\sigma = \frac{\mu Y s_1^T + \lambda Y s_1^N}{n_{E1} + n_{E1}\frac{(\psi\tau^{1-\sigma}-1)}{(\tau^{1-\sigma}-\psi)}\tau^{1-\sigma}} + \frac{\left(\mu Y s_2^T + \lambda Y s_2^N\right)\tau^{-\sigma}}{n_{E1}\tau^{1-\sigma} + n_{E1}\frac{(\psi\tau^{1-\sigma}-1)}{(\tau^{1-\sigma}-\psi)}}$$

$$n_{E1}^* = \frac{\mu Y s_1^T + \lambda Y s_1^N}{F\sigma\left(1 + \frac{(\psi\tau^{1-\sigma}-1)}{(\tau^{1-\sigma}-\psi)}\tau^{1-\sigma}\right)} + \frac{\left(\mu Y s_2^T + \lambda Y s_2^N\right)\tau^{-\sigma}}{F\sigma\left(\tau^{1-\sigma} + \frac{(\psi\tau^{1-\sigma}-1)}{(\tau^{1-\sigma}-\psi)}\right)}$$

An equivalent solution method can be used to solve for the optimal number of service sector firms, n_{S1}^* and n_{S2}^*.

System of Equations for 3-Region Model

$$F_1(n_{E1}, n_{S1})\sigma = \frac{\mu Y s_1^T + \lambda Y s_1^N}{n_{E1} + n_{E2}\tau^{1-\sigma} + n_{E3}\tau^{1-\sigma}} + \frac{\left(\mu Y s_2^T + \lambda Y s_2^N\right)\tau^{-\sigma}}{n_{E1}\tau^{1-\sigma} + n_{E2} + n_{E3}\tau^{1-\sigma}} + \frac{\left(\mu Y s_3^T + \lambda Y s_3^N\right)\tau^{-\sigma}}{n_{E1}\tau^{1-\sigma} + n_{E2}\tau^{1-\sigma} + n_{E3}}$$

$$F_2(n_{E2}, n_{S2})\sigma = \frac{\left(\mu Y s_1^T + \lambda Y s_1^N\right)\tau^{1-\sigma}}{n_{E1} + n_{E2}\tau^{1-\sigma} + n_{E3}\tau^{1-\sigma}} + \frac{\mu Y s_2^T + \lambda Y s_2^N}{n_{E1}\tau^{1-\sigma} + n_{E2} + n_{E3}\tau^{1-\sigma}} + \frac{\left(\mu Y s_3^T + \lambda Y s_3^N\right)\tau^{-\sigma}}{n_{E1}\tau^{1-\sigma} + n_{E2}\tau^{1-\sigma} + n_{E3}}$$

$$F_3(n_{E3}, n_{S3})\sigma = \frac{\left(\mu Y s_1^T + \lambda Y s_1^N\right)\tau^{1-\sigma}}{n_{E1} + n_{E2}\tau^{1-\sigma} + n_{E3}\tau^{1-\sigma}} + \frac{\left(\mu Y s_2^T + \lambda Y s_2^N\right)\tau^{-\sigma}}{n_{E1}\tau^{1-\sigma} + n_{E2} + n_{E3}\tau^{1-\sigma}} + \frac{\mu Y s_3^T + \lambda Y s_3^N}{n_{E1}\tau^{1-\sigma} + n_{E2}\tau^{1-\sigma} + n_{E3}}$$

$$F_1(n_{E1}, n_{S1})\sigma = \frac{(1-\mu)Y s_1^T + (1-\lambda)Y s_1^N}{n_{S1} + n_{S2}\tau^{1-\sigma} + n_{S3}\tau^{1-\sigma}} + \frac{\left((1-\mu)Y s_2^T + (1-\lambda)Y s_2^N\right)\tau^{-\sigma}}{n_{S1}\tau^{1-\sigma} + n_{S2} + n_{S3}\tau^{1-\sigma}} + \frac{\left((1-\mu)Y s_3^T + (1-\lambda)Y s_3^N\right)\tau^{-\sigma}}{n_{S1}\tau^{1-\sigma} + n_{S2}\tau^{1-\sigma} + n_{S3}}$$

$$F_2(n_{E2}, n_{S2})\sigma = \frac{\left((1-\mu)Y s_1^T + (1-\lambda)Y s_1^N\right)\tau^{1-\sigma}}{n_{S1} + n_{S2}\tau^{1-\sigma} + n_{S3}\tau^{1-\sigma}} + \frac{(1-\mu)Y s_2^T + (1-\lambda)Y s_2^N}{n_{S1}\tau^{1-\sigma} + n_{S2} + n_{S3}\tau^{1-\sigma}} + \frac{\left((1-\mu)Y s_3^T + (1-\lambda)Y s_3^N\right)\tau^{-\sigma}}{n_{S1}\tau^{1-\sigma} + n_{S2}\tau^{1-\sigma} + n_{S3}}$$

$$F_3(n_{E3}, n_{S3})\sigma = \frac{\left((1-\mu)Y s_1^T + (1-\lambda)Y s_1^N\right)\tau^{1-\sigma}}{n_{S1} + n_{S2}\tau^{1-\sigma} + n_{S3}\tau^{1-\sigma}} + \frac{\left((1-\mu)Y s_2^T + (1-\lambda)Y s_2^N\right)\tau^{-\sigma}}{n_{S1}\tau^{1-\sigma} + n_{S2} + n_{S3}\tau^{1-\sigma}} + \frac{(1-\mu)Y s_3^T + (1-\lambda)Y s_3^N}{n_{S1}\tau^{1-\sigma} + n_{S2}\tau^{1-\sigma} + n_{S3}}$$

Table 5

2-digit NASIC	Description	6-digit NAICS	Classification
11	Agriculture, Forestry, Fishing and Hunting		Neither
21	Mining		Neither
22	Utilities		Neither
23	Construction	236117, 236118, 237110, 237120, 238160, 238170, 238190, 238210, 238220, 238310, 238320, 238330, 238340, 238350, 238390, 238990	Service
31-33	Manufacturing	312120, 312130, 312140	Entertainment
42	Wholesale Trade		Neither
44-45	Retail Trade	445120, 445291, 445292, 445299, 445310, 446110, 446120, 446130, 446191, 446199, 447110, 447190, 448110, 448120, 448130, 448140, 448150, 448190, 448210, 448310, 448320, 451110, 451120, 451130, 451140, 451211, 451212, 451220, 452111, 452112, 452910, 452990, 453110, 453210, 453220, 453310, 453910, 453920, 453930, 453991, 453998, 454210, 454310, 454311, 454312, 454319, 454390	Entertainment
		441110, 441120, 441210, 441221, 441222, 441228, 441229, 441310, 441320, 442110, 442210, 442291, 442299, 443111, 443112, 443120, 443130, 443141, 443142, 444110, 444120, 444130, 444190, 444210, 444220, 445110, 445120, 445210, 445220, 445230, 454310, 454311, 454312, 454319	Service
48-49	Transportation and Warehousing	487110, 487210, 487990	Entertainment
51	Information	512120, 512131, 512132, 519120	Entertainment
52	Finance and Insurance	522120, 522130, 522190, 522210, 522220, 522291, 522292, 523910, 523930	Service
53	Real Estate Rental and Leasing	532292, 532299	Entertainment
		531110, 531130, 531190, 531210, 531311, 531312, 531320, 532111, 532120, 532210, 532220, 532230, 532291	Service

Note: Any codes not listed in the table are classified as "Neither."

2-digit NASIC	Description	6-digit NAICS	Classification
54	Professional, Scientific, and Technical Services	541110, 541120, 541211, 541213, 541921, 541940	Service
55	Management of Companies and Enterprises		Neither
56	Administrative and Support and Waste Management and Remediation Services	561510, 561520	Entertainment
		561622, 561710, 561740, 561790	Service
61	Educational Services	611110, 611210, 611410, 611420, 611430, 611511, 611512, 611513, 611519, 611610, 611620, 611630, 611691, 611692, 611699	Service
62	Health Care and Social Assistance	621111, 621112, 621210, 621310, 621320, 621330, 621340, 621391, 621399, 621410, 621491, 621492, 621498, 621511, 621512, 621610, 621910, 621999, 622110, 622210, 622310, 623110, 623210, 623311, 623312 ,623990, 624110, 624120, 624190, 624210, 624221, 624229, 624310, 624410	Service
71	Arts, Entertainment, and Recreation	711110, 711120, 711130, 711190, 711211, 711212, 711219, 711510, 712110, 712120, 712130, 712190, 713110, 713120, 713210, 713290, 713910, 713920, 713930, 713950, 713990	Entertainment
		713940	Service
72	Accommodation and Food Services	721110, 721120, 721191, 721199, 721211, 721214, 721310, 722110, 722211, 722212, 722213, 722310, 722320, 722330, 722410, 722511, 722513, 722514, 722515	Entertainment
81	Other Services (except Public Administration)	812199, 812990	Entertainment
		811111, 811112, 811113, 811118, 811121, 811122, 811191, 811192, 811198, 811211, 811411, 811412, 811420, 811430, 811490, 812111, 812112, 812113, 812191, 812199, 812210, 812220, 812310, 812320, 812910, 812922, 812930, 812990, 813110, 813211, 813212, 813219, 813311, 813312, 813319, 813410	Service
92	Public Administration		Neither

Note: Any codes not listed in the table are classified as "Neither."

Table 6

# of Reviews (s_{jt})	(1) OLS	(2) FE
# of Service Firms (n_{Sjt})	-0.9620***	-2.3556***
	(0.2678)	(0.9008)
# of Entertainment Firms (n_{Ejt})	2.7863***	5.2002***
	(0.5164)	(1.4708)
Y_{50}	-0.2466	3.38856**
	(0.5924)	(1.5901)
# of Occupied Houses	0.8169*	2.5091
	(0.4904)	(2.0784)
PUMA FE	No	Yes
Year FE	Yes	Yes
Adjusted R^2	0.3420	0.3671
F	13.72	12.53
N Obs	2974	2974

References

Alyakoob, M. and Rahman, M. S. (2018). Shared prosperity (of lack thereof) in the sharing economy.

Anas, A. and Xu, R. (1999). Congestion, land use, and job dispersion: A general equilibrium model. *Journal of Urban Economics*, 45.

Arellano, M. and Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *The Review of Economic Studies*, 58(2).

Arellano, M. and Bover, O. (1995). Another look at the instrumental variable estiamtion of error-component models. *Journal of Econometrics*, 68(1).

Blundell, R. and Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87(1).

Brousseau, F., Metcalf, J., and Yu, M. (2015). *Analysis of the impact of short-term rentals on housing*. Technical report, City and County of San Francisco, San Francisco, CA.

Cócola-Grant, A. (2016). Holiday rentals: The new gentrification battlefront. *Sociological Research Online*, 21(3).

Cócola-Grant, A. (2018). *Struggling with the leisure class: Tourism, gentrification and displacement*. PhD thesis, Cardiff University.

Cohen, J. (2018). Boston moves to regulate short-term rentals like Airbnb. https://nextcity.org/daily/entry/boston-moves-to-regulate-airbnb.

Fujita, M. (1988). A monopolistic competition model of spatial agglomeration: Differentiated product approach. *Regional Science and Urban Economics*, 18(1).

Fujita, M., Krugman, P., and Venables, A. J. (2001). *The Spatial Economy*. MIT Press.

Getz, D. (1993). Planning for tourism business districts. *Annals of Tourism Research*, 20.

Gitelson, R. J. and Crompton, J. L. (1983). The planning horizons and sources of information used by pleasure vacationers. *Journal of Travel Research*, 21(3).

Gotham, K. F. (2005). Tourism gentrification: The case of New Orleans' Vieux Carra (French Quarter). *Urban Studies*, 42(7).

Graviari-Barbas, M. and Guinand, S., editors (2017). *Tourism and Gentrification in Contemporary Metropolises*. Routledge.

Gurran, N. and Phibbs, P. (2017). When tourists move in: How should urban planners respond to Airbnb? *Journal of the American Planning Association*, 83(1).

Guttentag, D. A. and Smith, S. L. (2017). Assessing Airbnb as a distruptive innovation relative to hotels: Substitution and comparative performance expectations. *International Journal of Hospitality Management*, 64.

Honan, K. (2018). New York city council passes bill to regulate Airbnb. https://www.wsj.com/articles/new-york-city-council-passes-bill-to-regulate-airbnb-1531952763.

Ioannides, D., Röslmaier, M., and van der Zee, E. (2018). Airbnb as an instigator of 'tourism bubble' expansion in Utrecht's Lombok neighbourhood. *Tourism Geographies*.

Judd, D. R. (1991). Promoting tourism in US cities. *Tourism Management*, 16(3).

Krugman, P. (1980). Scale economies, product differentiation, and the pattern of trade. *The American Economic Review*, 70(5).

Lee, D. (2016). How Airbnb short-term rentals exacerbate Los Angeles's affordable housing crisis: Analysis and policy recommendations. *Havard Law and Policy Review*, 10.

Loudenback, T. (2018). Los Angeles just took a big step toward regulating airbnb stays — here's what new short-term rental laws could mean for hosts and travelers. https://www.businessinsider.com/airbnb-vrbo-regulations-los-angeles-what-it-means-for-hosts-renters-2018-5.

New York State Attorney General (2014). *Airbnb in the City*. Technical report, New York State Attorney General, New York, NY.

Roodman, D. (2009). A note on the theme of too many instruments. *Oxford Bulletin of Economics and Statistics*, 71(1).

Sans, A. A. and Domínguez, A. Q. (2016). Unravelling airbnb: Urban perspectives from Barcelona. In Russo, A. P. and Richards, G., editors, *Reinventing the Local in Tourism: Producing, Consuming and Negotiating Place*. Channel View.

Schäfer, P. and Braun, N. (2016). Misuse through short-term rentals on Berlin housing market. *International Journal of Housing Markets and Analysis*, 9(2).

Shoval, N., McKercher, B., Ng, E., and Birenboim, A. (2011). Hotel location and tourist activity in cities. *Annals of Tourism Research*, 38(4).

Sivitanidou, R. and Wheaton, W. C. (1992). Wage and rent capitalization in the commercial real estate market. *Journal of Urban Economics*, 31.

Versichele, M., de Groote, L., Bouuaert, M. C., Neutens, T., Moerman, I., and de Weghe, N. V. (2014). Pattern mining in tourist attraction visits through association rule learning on Bluetooth tracking data: A case study of Ghent, Belgium. *Tourism Management*, 44.

Wooldridge, J. (2010). *Econometric Analysis of Cross Section and Panel Data*. MIT Press.

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).