Who Gets to Share in the “Sharing Economy”: Understanding the Patterns of Participation and Exchange in Airbnb[1]

Mehmet Cansoy, Boston College

&

Juliet Schor, Boston College

May 2016

Abstract:

Recently, Airbnb has released a number of studies that claim that the company is having a positive impact in the regions it operates in by driving tourists and revenue to areas underserved by traditional hotels, providing lower-income residents with a revenue stream and building up interpersonal trust (Airbnb, 2015). However, these claims have yet to be validated by independent research and they are further complicated by recent findings showing that racial discrimination against guests is a real concern in the platform (BG Edelman, Luca, & Dan, 2015). Issues of inequality in access and use go beyond the Airbnb platform and have been a consistent concern in both popular and academic literature on the “sharing economy.” Yet we still do not fully understand how these new institutions interact with the existing socio-economic cleavages, at least partly because large scale data on the “sharing economy” has not been easily available. In this paper, we overcome this lack of data using a unique dataset containing information about roughly 125,000 Airbnb listings from 104 metropolitan statistical areas in the US. We match the information found on the Airbnb platform to the US Census data on the census-tracts in which listings are located. This allows us to study how income, race and education impact the economic activity on the platform. We show that while the platform appears to favor lower income and more racially diverse areas at first glance, the true patterns of participation and exchange are much more complicated. Despite the fact that census tracts with higher income and a greater proportion of white people have less listings and fewer reviews, these still have better ratings and demand higher prices. More importantly, we show that education is a heavily influential factor in all of these patterns, suggesting that these new institutions might be benefitting individuals with higher cultural capital disproportionately. While our data has some inherent limitations, because we only study the hosts on the platform and do not have complete information on the exchanges, we nonetheless believe that our findings are a significant contribution to the debate on the sharing economy and its future.

---

[1] Preliminary draft, please do not cite or disseminate without contacting the authors.

Introduction:

A recent report, publicized by Airbnb proclaimed that Airbnb offers "middle-class families facing the squeeze of stagnant incomes help to make up the difference" (Sperling, 2015). Similar claims about the online accommodation sharing company are abundant in various media, ranging from guides on how to maximize earnings (Kessler, 2013) to discussions on the social benefits of the service (Lott, 2014). However, coverage about Airbnb, like many other "sharing economy" companies, has not been all positive. Criticism of the platform has mostly been focused on units that are exclusively being used as a rental properties while bypassing regulations and taxes (Streitfeld, 2014), with several local authorities taking action or threatening to do so (Berton & Wecker, 2015). In response Airbnb has released a number of studies that claim that the company is having a positive impact in the regions it operates in by driving tourists and revenue to areas underserved by traditional hotels, providing lower-income residents with a revenue stream and building up interpersonal trust (Airbnb, 2015).

These claims have yet to be validated by independent research and they are further complicated by recent findings showing that racial discrimination against guests is a real concern in the platform (BG Edelman et al., 2015). The biggest impediment to answering these complex questions about the patterns of participation in the sharing economy and their consequences has been the lack of data. In this paper, using a unique dataset containing information about roughly 125,000 Airbnb listings from 104 metropolitan statistical areas in the US[2], we attempt to map the patterns of participation and exchange on Airbnb. We match the information found on the Airbnb platform to the census-tracts in which listings are located, and show that while the platform appears to favor lower income and more racially diverse areas at first glance, the true patterns are much more complicated. Despite the fact that census tracts with higher income and a greater proportion of white people have less listings and fewer reviews, these still have better ratings and generate more revenue. More importantly, we show that education is a heavily influential factor in all of these patterns, driving participation, exchange and revenue.

Literature Review:

Airbnb is a web-based marketplace in which people can list their accommodations for (usually) short-term rental. The listings range between a futon in a studio apartment to entire medieval castles in Europe. The company, founded in 2007, makes money by charging a commission on any rentals that take place on their platform. Currently it is active worldwide, and has more than 2,000,000 unique accommodation options listed for rent (Airbnb, 2016). However most of what we know about its functioning and impact is based on the company's own reports.

These reports focus on the experiences of hosts, guests and the impact of the company on twelve major cities around the world (Airbnb, 2015). The summary findings on hosts highlight the fact that about half of all hosts are low to moderate income, and report using the income they make through the platform for essential expenses. On the guest side, the reports highlight the fact that guests prefer to experience their destinations "like a local" and explore specific neighborhoods,

---

[2] The results presented here are based on preliminary analysis, we are currently still adding listings to the data and improving our analytical methods, as a result of which the findings might change. The details of how the data was collected, cleaned and analyzed are below.

spending more than double the time and money compared to the average visitor. This goes hand-in-hand with their findings that almost three quarters of the listings are outside “the main hotel districts” and about half of the guest spending stays in these areas. However, the reports are very brief, the findings not elaborated, and the methods and data they use is not presented in any detail. Perhaps more importantly, some important aspects of their findings are contradicted by other sources.

For example, there are serious claims that “professional hosts” with two or more properties on the site, account for a very large portion of the revenue (Kang, 2016). In New York, these claims have come to a head, with the Attorney-General of the state arguing that these practices are illegal (Streitfeld, 2014). The company was forced to share its user data in November 2015, after a lengthy court battle (Isaac & Mike, 2015). However, there is evidence to suspect that the data might have been manipulated by a purge of the worst offenders from the Airbnb platform in the weeks leading up to the data release (Cox & Slee, 2016).

While we partially adress this controversy in the data section below, we believe that the debate over professional hosts should not take our focus away from the larger question of who gets to benefit from the opportunity presented by Airbnb. There is a wide-ranging “neighborhood effects” literature on the inequalities produced and re-produced by housing patterns, and Airbnb has the potential to be influencing these dynamics because it allows people to monetize their housing. Researchers have focused on crime (Krivo & Peterson, 2000; M. R. Lee, 2000), poverty (Clampet-Lundquist & Massey, 2008; Ludwig et al., 2008; Parisi, Lichter, & Taquino, 2011; Quillian, 2012; Sampson, 2008; Swisher, Kuhl, & Chavez, 2013) and psychological well-being (Ross, Reynolds, & Geis, 2000) in conjunction with race and racial segregation to understand how housing patterns impact a number of social outcomes. In this literature the causal mechanism through which housing patterns impact these outcomes pivots around the social ties within neighborhoods. These ties are theorized to provide and restrict access to material and psychological resources as well as economic opportunities.

While a similar social connection might be driving Airbnb listings –with membership and utilization of the platform spreading through the community with neighbors imitating one another- we are much more interested in the patterns of participation in Airbnb and what they can reveal about social cleavages. There is reason to expect that participation of hosts in Airbnb might be predicated upon factors such as race, class and education. First, participation not only requires access to capital (having a place to rent, living in a location desireable for travelers, furnishing it, maintaining an online presence) but also requires the cultural competency to navigate the often complex dynamics of making and managing bookings. Thus the location of individuals across the cleavages of class, race, and education has a relatively direct relationship with their participation. Perhaps more interestingly, there are theoretical reasons to focus on these cleavages as well.

The relationship between class (or at least income) and sharing is theorized in a relatively straightforward manner in the field of economics. Individuals are understood to enage in sharing so long as the marginal utility they draw from it is greater than the marginal costs it imposes on them. Thus, in simplified terms, sharing is expected to be more prevalent among the poor, where the marginal benefits it brings (access to goods and services) are high than among the rich where they are lower (since they already have access through ownership or through the market). Recently, Salcedo, Schoellman, and Terlitt (2012) have used this approach to explain the reduction in household sizes between 1850 and 2000, as an outcome of income growth. However, Benkler (2004) has argued that decentralized institutions for sharing (like Airbnb) lower the costs of sharing for participants, and can thus facilitate greater participation across the income divide. While isolating the impact of income alone on the propensity to engage in sharing remains a challenge (because the perceived costs and utility of engaging in sharing is dependent on norms and preferences), preliminary work by Fremstad suggests that Benkler's vision for these new institutions might not be accurate. Using polling data, in conjunction with the census, he shows that people with lower income participate in these institutions at a higher rate compared to those with higher incomes. However, his work focuses more on the consumption end of the process and it is not immediately clear if the production side is subject to the same dynamics. His findings are also complicated by Thebault-Spieker, Terveen and Hecht (2015). Focusing on Chicago, they show how workers on the TaskRabbit platform, which allows people to hire workers for specific one-off tasks, are concentrated in relatively well-off urban areas and are unwilling to travel to lower socioeconomic-status areas for jobs. They argue that this results in the the poorer areas being shut-off from the potential benefits of the sharing economy.

On the other hand, Edelman, Luca and Dan (2015) have found, through an audit study, that guests with distinctively African-American names are less likely than those with distinctively White names to make a successful reservation through the Airbnb platform. This effect holds true across hosts of different races and hosts engage in this discriminatory practice even though in many cases this means lost revenue for them. In an earlier study Edelman and Luca (Benjamin Edelman & Luca, 2014) have also found that listings with a non-black host charged 12% more per night than comparable listings with a black host. Taken together, these findings suggest that race is an important dynamic that influences the patterns of participation and exchange in the sharing economy.

We also believe that education is also a key factor in explaining these patterns. Previous research suggests that highly educated people participate in the sharing economy at a greater rate and subcribe to an ethos of sharing that values their engagement in it beyond that of a simple utilitarian logic.

Methods:

In order to understand the patterns of opportunity, and investigate potential exclusion along class and race lines, in the Airbnb platform, we chose to analyze the listings matched with US Census data at the census tract level. This decision was based on three reasons. The first, and for us the most important, one is that this was a way to protect potentially sensitive information of Airbnb users, while still contributing to the debate over sharing economy and its opportunity structure in a meaningful way. We take the debate over the ethics of using online data (Zimmer, 2010) seriously, and while we believe that in our case the data is more unambiguously public than in other cases[3], we also think that using tract-level data was adequate to conduct our analysis. Second, we chose the level of the census tracts as our unit of analysis because a much broader range of data is available on them compared to the smaller census blocks. This allows us to test a number of variables that might be influencing the participation structure for Airbnb hosts, while still retaining a relatively small geographical unit. Moreover, while there are some advantages to using census blocks, especially in an HLM context (Parisi et al., 2011, p. 835), both of the current geographical studies on the sharing economy (BG Edelman et al., 2015; Thebault-Spieker et al., 2015) have used census tracts and we believe it is important for our results to be comparable to theirs. This level of aggregation (often as a proxy for a "neighborhood") has also been used by many researchers in the neighborhood effects literature (Krivo & Peterson, 2000; R. Lee et al., 2009; Quillian, 2012) We use two types of mixed-effects models based on the distribution of our dependent variables, which we explain in further detail below.

Data:

The information on Airbnb listings was collected by an automated program which queried the website's search function with geographical locations.[4] These locations were city and neighborhood names in metropolitan statistical areas (MSAs) in the USA which have a population above 500,000.[5] Each location name was queried on four separate occasions between November 2015 and April 2016. All of the unique listings returned by the query were then processed to extract their geographical location (latitude and longitude), alongside other data which we present below. The geographical location was then matched[6] to individual census tracts using the US Census' Geocoder API (US Census 2016c). We then merged the listing-level data with the 5-year estimates of the American Community Survey's 2014 Results (US Census Bureau 2016d) for the same census tract.

This process provided us with information on roughly 130,000 Airbnb listings in the relevant MSAs. We are currently still processing and cleaning some this data, and therefore the results presented below are preliminary and subject to change. While we have no certain way of knowing how many of the actual Airbnb listings we were able to capture with our methods, comparison to

[3] The listings are publicly accessible (without requiring an account) on the Airbnb platform and for many major cities scraped Airbnb listing data, including some individual information about the hosts, is provided by Inside Airbnb (www.insideairbnb.com).

[4] The program is broadly similar to the one described by Tom Slee (2016) and the code provided by Hamel Husain (2016). We would like to extend our thanks to both of them for sharing their code and coding logic.

[5] The metropolitan area population statistics are based on 2014 Annual Estimates of the Resident Population (US Census Bureau, 2016a). There are 104 MSAs within the USA (excluding non-state territories) that fit this criteria and were included in our study, see Appendix 1 for a list. The names of cities in the MSAs are from the table "Principal cities of metropolitan and micropolitan statistical areas" from the US Census Bureau (US Census Bureau, 2016b). There are 224 principal cities in the MSAs that fit our criteria. The neighborhood names within the cities are from "Zillow Neighborhood Boundaries" data. There are 4629 neighborhoods that are located within the cities which fit our criteria. A list of cities and neighborhood names is available upon request.

[6] Airbnb does not provide accurate geographical location of listings, but provides a roughly 0.5 mile-wide "circle" within which the property is located. We used the center of these circles as the locations of the listings. While it is possible that this might result in faulty matches to census tracts, we believe that given the limited nature of the data and the size of Census Tracts, the level of aggregation is appropriate and that any mismatches will be randomly distributed across tracts.

another publicly available dataset on Airbnb listings for the same period suggests that we have comprehensive coverage outside of New York City.

(Figure 1 about here)

In Figure 1, you can see that in our sample, the "shared room" listings, in which the host and the guest are expected to sleep in the same room, make up a negligible 3% of the sample. Of the other two types of listings, "Entire home/apartment" listings in which the guest rents the whole location without the host being present during the rental, make up about 56% of the sample and the "private room" listings in which the guest has access to a private room the remaining 40%. Interestingly, and in contrast with Airbnb's reports, we find that almost 42% of all listings are posted by hosts with 2 or more listings. In Figures 2 and 3 below, you can see that of the three types of listings, entire home/apartment listings are the most expensive per night and per night per person, followed by private rooms and shared rooms. These figures also show that listings managed by hosts that have 2 or more listings tend to be cheaper than listings managed by hosts with a single listing.

(Figure 2 and 3 about here)

Finally in Figure 4 you can see that the listings available to be booked instantly (without needing the host's approval) are unequally distributed among the two types of hosts. Among listings managed by hosts with a single listings, the instant booking listings make up about 14% of all available listings, whereas among listings managed by hosts with 2 or more listings the figure is more than 21%.

(Figure 4 about here)

In our analysis we use a number of variables based on the data obtained through the Airbnb pages of listings, which we explain in detail below.

Number of Listings: This variable measures the total number of listings in a census tract. Census tracts with no listings were assigned a value of zero. Since this is a count variable with a large number of zero values, we used a mixed-effects negative binomial model in our analysis, with the census tracts clustered at the MSA level. Exposure was assumed to be uniform across all census tracts and thus was not modeled, because there is no reliable way to trace Airbnb's roll-out in the various MSA's, and the platform has been available nation-wide for a number of years.

Nightly Price: This is the nightly price of a listing as it was listed at the most recent date the listing was accessed. This variable was top- and bottom-coded so that the lowest and highest values would be 10$ and 8000$, respectively. Due to deviations from normality, the variable was then log transformed. We used a mixed-effects model in our analysis of this variable, with listings clustered at the census-tract level.[7]

[7] 3-level mixed effects models we ran, with tracts clustered at the level of MSAs required us to drop roughly 6000 observations that were the only observations within a census tract because they made the model unidentified. Furthermore, the resulting models had lower intra-class correlation and substantively similar results. Therefore we opted to present the simpler 2-level models.

Nightly Price per Person: This is the nightly price of a listing (see above) divided by the number of people that the listing can accommodate (top-coded to 30 by Airbnb). Due to deviations from normality, the variable was then log transformed. We used a mixed-effects model in our analysis of this variable, with listings clustered at the census-tract level.[8]

Number of Reviews: On Airbnb, guests are encouraged to provide a review of listings they have stayed in. This variable codes the number of such reviews for a listing that is visible to a guest (the hosts can see some reviews before they are made public). Since this is a count variable with a large number of zero values, we used a mixed-effects negative binomial model in our analysis, with listings clustered at the census-tract level. Exposure was modeled by multiplying by the number of months the host has been a member of the platform (see below) by 30 and then dividing the resulting number by the length of the minimum stay required for the listing. We are aware that this is not an ideal exposure measure. In addition to problems associated with the host's membership date (see below) the length of the minimum stay for a given property can change throughout the year, and can be changed readily by the host. However, in the absence of better data, we still believe that it is an adequate way to model exposure to potential reviews.

'Reverse' Rating: On Airbnb, guests also provide a numerical rating for listings they have stayed in, on a scale ranging from 0 to 10 (on the site this is displayed as a 5-star scale with half stars). While the guests can rate several aspects of a listing like location or check-in process separately we only focused on the overall rating and captured the average of all ratings on a scale from 0-100. However, bad ratings on the site are very rare and in our sample more than 31% of all listings with a rating had a perfect rating of 100 and a further 60% had ratings between 85 and 99. In order to analyze the values we reverse-coded the variable so that a listing with a 100 rating would have a value of 0 and higher scores would indicate worse listings. This netted us a distribution that closely approximated a negative binomial distribution, even though the variable was not a count variable. So we used a mixed effects negative-binomial model in our analysis. Exposure was assumed to be uniform across all listings and thus was not modeled, because ratings only become publicly visible on the platform after a certain number of reviews have been left for a listing.[9]

Imputed Income: For every listing in our dataset, we calculated an estimate of the income they have generated on Airbnb. This was done by multiplying the nightly price for a listing by its minimum stay requirement and the total number of reviews. This method of estimating income is likely to result in significantly lower values than the actual income for listings, since guests can stay longer than the minimum stay requirements, pay fees not included in the price to the host and might not review the listing after their stay. Additionally, the price for the listing might have been changed over the period it has been on Airbnb, and assuming that the current price is comparable to past prices is clearly problematic. However, in the absence of actual transaction data from

[8] 3-level mixed effects models we ran, with tracts clustered at the level of MSAs required us to drop roughly 6000 observations that were the only observations within a census tract because they made the model unidentified. Even after this the negative binomial models were non-convergent. The linear models, on the other hand, had lower intra-class correlation and substantively similar results. Therefore we have opted to present the simpler 2-level models.

[9] Our data suggest that the cut-off is 2 reviews, but we do have a small number of listings with more than 2 reviews that did not have a rating.

Airbnb we believe this is the best estimate for listing incomes. The variable was only calculated for listings with at least 1 review and top-coded so that the highest values would be 100,000$. It was then log-transformed due to deviations from normality.

Length of Membership: For every listing we calculated how many months the user had been a member of Airbnb. This is clearly not a perfect way to measure the length of time a given listing has been on the platform, since hosts do manage multiple listings (at the same time and concurrently). However, it is still the best measure available and we use it as a control variable in many of our models, as well as a way to model exposure to potential number of reviews (see above). The variable is standardized and grand mean centered for ease of interpretation.

Number of Amenities: For every listing we calculated a variable that indicated the number of amenities the listing had or provided to guests. The potential amenities (the most widely available amenities other than the safety equipment) includes A/C, Breakfast, Cable TV subscription, doorman, dryer, essentials, fireplace, gym, hairdryer, hangers, hot tub, intercom, internet access, clothes iron, kitchen access, on-site parking, pool, shampoo, TV and washing machine. This is intended to capture the quality of the listing and is used as a control variable in many of our models. The variable is standardized and grand mean centered for ease of interpretation.

In our analysis we also use a number of variables drawn from the 5-year estimates of the 2014 American Community Survey (U.S. Census Bureau 2016d). All of these variables are standardized and grand mean centered for ease of interpretation.

Total Population of MSA: This is the total population for the MSA that a given listing or tract is in.

Total Population of the Census Tract: This is the total population of a given census tract.

Median Income: This is the median income for all households within a census tract.

% with BA: This variable is the percentage of the total population 25 and over in a census tract that has a BA degree or higher educational credentials.

% non-White: This variable is the percentage of the total population in a tract that did not identify as White (including those that identified as more than one race, even if one of the races was White).

% Black: This variable is the percentage of the total population in a tract that identified as Black (excluding those that identified as more than one race, even if one of the races was Black).

% Hispanic: This variable is the percentage of the total population in a tract that identified as Hispanic (excluding those that identified as more than one race, even if one of the races was Hispanic).

The descriptive statistics for all of these variables can be found in Appendix 2. We chose to handle missing data (from the US Census data) with listwise deletion to keep our models simpler, since the number of cases for which there was any missing data is at most around 1% of all valid cases we believe that listwise deletion does not impact our results in any substantial manner.

Limitations:

The nature of the data we are using places a number of significant limitations on our analysis. First of all, despite being as thorough as possible, we cannot be sure that we captured all of the listings within the geographies we are interested in. Despite comparable results to other Airbnb scraping efforts, we simply have no way of establishing our coverage rate, short of obtaining data from Airbnb. However we still believe that this is the best way to study Airbnb listings and data scraping can potentially answer the dearth of large-scale data in the studies of the sharing economy.

The second limitation of our data is that it is cross-sectional. Despite searching for listings in the same location four times, the time differences between the searches we conducted were not meaningful, and we chose to aggregate results from the four searches. This means that we are not able to control for time and time-variant factors in our analysis. All of our results at this stage should be interpreted as correlational, rather than causal.

The third and most important limitation of our data is that hosts self-select into participating in Airbnb and since we study the listings that are already on the platform and without access to individual-level data, we have no way of identifying trends at this level that might be influencing the patterns of participation we study. While this means that the results of our analysis might not be capturing all of the factors that create these patterns, we still believe that short of obtaining data from Airbnb and consent from hosts, studying neighborhood-level patterns is the most ethical and efficient approach to this type of research.

Results:

a. Number of Listings:

(Table 1 about here)

The results of our analysis on the number of listings in a census tract, with the tracts clustered at the MSA-level can be found in Table 1 above. The first thing to note here is that our variables of interest, in models 2 and 3, improve model fit significantly when compared with model 1 with only the control variables. We can also see that the census tract population has an incidence rate ratio that is consistently higher than 1 across all three models, while MSA population has one that is below 1 in them. This indicates that census tracts with higher population have more listings, all else being equal, whereas census tracts in MSAs with higher population have less listings than comparable census tracts in other MSAs. The former findings is relatively easy to explain, since highly populated tracts will have more accommodations. The latter, on the other hand, is likely explained by the larger geographic size of the more populous MSAs, which results in them having many census tracts with few or no listings.[10]

[10] We also know that in the case of New York our scraping is currently underperforming, and the size of New York and the number of listings in there might be biasing our results in this case.

We also find that census tracts with higher median incomes have fewer listings. In fact our data suggests that census tracts that are a standard deviation above the mean median income level are likely to have roughly 43% of the number of listings in a census tract at the mean. Race seems to play a similarly important role in hosts' participation on Airbnb. One standard deviation increase in the non-white population is correlated with almost a 60% increase in the number of listings. However, when we study race in finer detail, we find that while both the percentage of the population that is Black or Hispanic have positive incidence rate ratios, the one for the percentage of Hispanic population is much higher. Education is by far the strongest predictor in the model, a standard deviation increase in education is associated with an almost 5-fold increase in the number of listings in a census tract.

b. **Nightly Price and Nightly Price Per Person**:

(Tables 2 and 3 about here)

The results of our analysis of nightly price and nightly price per person of Airbnb listings can be found in Tables 2 and 3 above. With both independent variables we can see that clustering explains 25-30% of total variance, and the inclusion of our independent variables of interest increases model fit considerably. In terms of control variables, we find that while listings whose hosts have been on the platform longer tend to have somewhat higher nightly prices, they are actually cheaper per person per night. The number of amenities follow a similar logic, in which a higher number of amenities are associated with higher nightly prices, but lower per person per night prices. This suggests that large properties with many amenities were among the earlier adopters of the platform, not a surprising find given the fact that some of the properties are vacation rentals or similar properties that were already being rented out and simply used the platform as a marketing tool. Census-tract population, is negatively associated with both independent variables, suggesting that listings in more populous tracts have lower prices. MSA population, on the other hand is not significant predictor of nightly price, when controlling for income race and education, but is positively associated with nightly price per person. This suggests that listings in highly populated urban areas charge higher prices.

The median income in a census tract is positively associated with both nightly price and nightly price per person. The percentage of population with a BA degree or higher also has positive coefficients when predicting both dependent variables. While the strength of its relationship with nightly price is comparable to that of income and race, when we look at price per person per night, we again see that education is by far the strongest predictor. The variables that measure race are overall negatively associated with both of the price variables. However, the percentage of population that is non-white is not a significant predictor of the per person per night price, and the coefficients for the percentage of the population that is Black or Hispanic are both significantly less powerful predictors of this variable then they are of nightly price.

c. Number of Reviews:

(Table 4 about here)

The results of our analysis of number of reviews per listing, clustered at the census-tract level can be found in Table 4 above. The goodness-of-fit measures show that the addition of the variables of interest to the model improve model fit substantially. In a pattern analogous to what we found for the price variables, we see that listings in more populous tracts have fewer reviews, while those in more populous MSAs have more. The number of amenities in a listing also has an incidence rate ratio above one, indicating that listings with more amenities have more reviews. Length of membership is not included in this model because it was used to calculate an exposure variable (see data section above).

We see that the incidence rate ratio for the income variable is lower than 1, indicating that listings in higher income tracts have fewer reviews. Education, following the same trends we have uncovered in the preceding sections, is associated with higher numbers of reviews, and is the strongest predictor. Perhaps the most important thing to note in this table is that it shows a relatively important divide between the impacts of the percentages of the population that is Black and Hispanic. While both have incidence rate ratios above 1, a standard deviation change in the % of the population that is Black is associated with only a 2.4% increase in the number of reviews, whereas a comparable increase in the Hispanic population is associated with an increase more than four times that.

d. Rating:

(Table 5 about here)

In Table 5 above you can see the results of our analysis of listings' ratings, clustered at the census tract level. It is important to remember that the dependent variable has been reverse coded so that higher values indicate worse ratings. The inclusion of the independent variables of interest increase model fit significantly in models 2 and 3, compared to model 1. The census tract level population variable is not significant in either of the two full models. On the other hand, the MSA level variable's significance denotes that listings in more populous MSAs tend to be rated worse. The number of amenities in a listing, as can be predicted, is associated with better ratings. Listings whose hosts are newer on the platform also tend to have better ratings – probably because the larger number of reviews they accrue increases the chances of getting a negative one or the willingness of guests to leave more negative reviews because they will not impact the host as much.

Education, which is a very strong predictor for all of the other dependent variables is not a significant predictor of ratings. Median income, on the other hand, has an incidence rate ratio lower than 1 in both of the full models, indicating that listings in higher income census tracts are rated better (since the dependent variable is reverse coded). Race, in all three measures, has the opposite relationship. Listings in racially diverse tracts tend to be rated worse, and the impact on ratings is comparable for the percentages of the population that are Black and Hispanic.

e. **Imputed Income:**

(Table 6 about here)

In Table 6 above, you can see the results of our analysis of the imputed income per listing, clustered at the census-tract level. We can see that clustering accounts for 14-18% of variance in the dependent variable and the inclusion of the independent variables of interest in the model improves model fit when comparing models 3 and 4 to the empty model and model 2. The length of membership of the host, as could be expected, is positively associated with imputed income, as is the number of amenities and MSA level population. Census tract population, on the other hand is negatively associated, suggesting that listings in more populous tracts have generated less revenue on the platform.

Income has a relatively weak but negative association with imputed income. Suggesting that listings in more affluent tracts have not made as much money as those in less affluent ones. This is in line with previous findings, since listings in these areas had lower numbers of reviews. Race, on the other hand, shows a complicated relationship with imputed income. The percentage of the population that is Hispanic is not a significant predictor, while the coefficient for the percentage of population that is Black is highly significant and negative. Finally, education is once again the strongest predictor in the model, with a positive coefficient.

Discussion:

Our results show that there is some merit to Airbnb's claims about the patterns of participation on their platform. Areas with lower median income and a higher proportion of non-white population have more listings and higher numbers of reviews per listing. This suggests that these areas might be benefitting as a result of their participation on the platform as tourists and their money is diverted into them.

However, we believe that our findings complicate Airbnb's claims in important ways. First of all, we find that listings in higher income census tracts have higher nightly prices and nightly prices per person. This is probably an indication of the rent they can extract from their real estate (since these tracts are more likely to be closer to travel attractions and business centers). However it is also a function of the Airbnb's decentralized institutional design, which allows hosts to set their own prices (as opposed to a structure like the ride-sharing platforms where the prices are set by the platform). While this might be increasing the participation of the higher income hosts, because it allows them to increase the "marginal utility" they gain from sharing, it is allowing them to capture a significant share of the exchange in the platform. Our findings show that listings in these affluent areas are also rated better and generate a greater amount of revenue than comparable areas with lower income. One can therefore argue that Airbnb favors these affluent areas in significant ways.

Our findings in terms of race also paint a similar picture. Census tracts in which the percentage of non-white population is higher, despite higher rates of participation, and higher numbers of reviews per listing, command lower prices, are rated worse, and generate less income. While this is in itself not proof of discrimination against hosts in these areas, it does show that participation in the sharing economy, and being able to benefit from such participation, is patterned by race. This is perhaps best seen in the stark differences between the impacts of the percentages of the population that are Black and Hispanic, where an increase in the latter is associated with greater participation, higher per person per night prices, more reviews, better ratings and greater revenue than an increase in the former.

Finally, we believe that education is the key variable that deserves a much greater focus in the context of the sharing economy. We find that it is consistently the strongest predictor of all of our dependent variables, excluding ratings. This has a a number of important implications. First, participation in the sharing economy is likely to be heavily norm-driven as education is such a strong predictor of the number of listings in a given tract. This is in line with our findings from other qualitative studies, which have found large numbers of highly educated individuals engaged in sharing economy activites across a large number of sites. Second, the finding that listings in census tracts with highly educated populations charge higher prices and get better ratings might be the result of homophily between guests and hosts where the guests are willing to pay a premium to stay with someone like themselves in a location where they feel at home. Finally, the fact that education is the only independent variable that is positively associated with imputed income is key to understanding the long-term impacts of Airbnb on the urban environment. If these trends hold, the extra rent that Airbnb provides could play a key role in gentrification, as people who don't participate on the platform at the same rate or with the same success get priced out of their neighborhoods.

In this paper, we have attempted to bring some important data to bear on some of the key debates around the sharing economy. Using scraped data from Airbnb in conjunction with US Census data on census tracts, we investigated the patterns of participation and exchange on the online accommodation platform, using mixed-effects models. Our focus was primarily on class, race and education and we found that while Aribnb's claims about helping the middle class, and diverting revenue to lower income and more diverse areas has some merit it needs to be qualified. High-income areas, despite participating on the platform at lower rates, still have listings that generate more income and have higher ratings. Despite high rates of participation and exchange, racially diverse areas usually have listings that are rated lower, and generate less income. Finally, education turns out to be the lynchpin to understanding the patterns of participation and exchange on the platform, as census tracts with more highly educated populations participate at higher rates, and have listings with higher prices, more reviews, better ratings and higher revenue.

References:

Airbnb. (2015). The Economic Impacts of Home Sharing in cities around the world. Retrieved March 2, 2015, from https://www.airbnb.com/economic-impact

Airbnb. (2016). About Us. Retrieved March 2, 2016 from https://www.airbnb.com/about/about-us

Benkler, Y. (2004). Sharing nicely: On shareable goods and the emergence of sharing as a modality of economic production. *Yale Law Journal*, *114*(2), 273–358.

Berton, E., & Wecker, K. (2015). Europe cracks down on Airbnb, other room-sharing sites. *USA Today*. Retrieved March 2, 2016 from http://www.usatoday.com/story/money/business/2015/07/06/europe-airbnb-room-sharing/29263881/

Clampet- Lundquist, S., & Massey, D. S. (2008). Neighborhood Effects on Economic Self-Sufficiency: A Reconsideration of the Moving to Opportunity Experiment. *American Journal of Sociology*, *114*(1), 107–143.

Cox, M., & Slee, T. (2016). *How Airbnb's data hid the facts in New York City*. Retrieved March 2, 2016 from http://insideairbnb.com/reports/how-airbnbs-data-hid-the-facts-in-new-york-city.pdf

Edelman, B., & Luca, M. (2014). Digital Discrimination: The Case of Airbnb.com. *Harvard Business School*, 21. http://doi.org/10.2139/ssrn.2377353

Edelman, B., Luca, M., & Dan, S. (2015). Racial Discrimination in the Sharing Economy: Evidence from a Field Experiment. *HBS Working Paper*, 1–34. Retrieved March 2, 2016 from http://www.benedelman.org/publications/airbnb-guest-discrimination-2015-12-09.pdf

Husain, H. (2016). Python Function To Scrape Airbnb. Retrieved September 5, 2015 from https://github.com/hamelsmu/AirbnbScrape

Isaac, & Mike. (2015, December 1). Airbnb Releases Trove of New York City Home-Sharing Data. *New York Times*, pp. 2–4. Retrieved March 2, 2016 from http://www.nytimes.com/2015/12/02/technology/airbnb-releases-trove-of-new-york-city-home-sharing-data.html

Kang, C. (2016, January 21). Airbnb Takes Its Case to U . S . Mayors Conference. *New York Times*, pp. 2–5. Retrieved March 2, 2016 from http://www.nytimes.com/2016/01/22/technology/airbnb-takes-its-case-to-us-mayors-conference.html

Kessler, S. (2013). Secrets Of Running A Six-Figure Airbnb Business. *Fast Company*. Retrieved March 2, 2016 from http://www.fastcompany.com/3021179/secrets-of-running-a-six-figure-airbnb-business

Krivo, L. J., & Peterson, R. D. (2000). The Structural Context of Homicide: Accounting for Racial Differences in Process. *American Sociological Review*, *65*(4), 547.

Lee, M. R. (2000). Concentrated poverty, race, and homicide. *Sociological Quarterly*, *41*(2), 189–206.

Lee, R., Castree, N., Lawson, V., Paasi, a., Radcliffe, S., & Withers, C. (2009). Progress in Human Geography? *Progress in Human Geography*, *33*(1), 3–6.

Lott, T. (2014, September 19). Airbnb : the social benefits of renting out a room. *The Guardian*. Retrieved March 2, 2016 from http://www.theguardian.com/lifeandstyle/2014/sep/19/airbnb-social-benefits-renting-out-room

Ludwig, J., Liebman, J. B., Kling, J. R., Duncan, G. J., Katz, L. F., Kessler, R. C., & Sanbonmatsu, L. (2008). What Can We Learn about Neighborhood Effects from the Moving to Opportunity Experiment? *American Journal of Sociology*, *114*(1), 144–188.

Parisi, D., Lichter, D. T., & Taquino, M. C. (2011). Multi-Scale Residential Segregation: Black Exceptionalism and America's Changing Color Line. *Social Forces*, *89*(3), 829–852.

Quillian, L. (2012). Segregation and Poverty Concentration: The Role of Three Segregations. *American Sociological Review*, *77*(3), 354–379.

Ross, C. E., Reynolds, J. R., & Geis, K. (2000). The Contingent Meaning of Neighborhood Stability for Residents' Psychological Well-Being. *American Sociological Review*, *64*(4), 581–597.

Salcedo, A., Schoellman, T., & Terlitt, M. (2012). Families as roommates: Changes in U.S. Household size from 1850 to 2000. *Quantitative Economics*, *3*, 133–175.

Sampson, R. J. (2008). Moving to Inequality: Neighborhood Effects and Experiments Meet Social Structure. *American Journal of Sociology*, *114*(1), 189–231.

Slee, T. (2016). Airbnb Data Collection: Methodology and Accuracy. Retrieved September 5, 2015 from http://tomslee.net/airbnb-data-collection-methodology-and-accuracy

Sperling, G. (2015). *How Airbnb Combats Middle Class Income Stagnation*. Retrieved March 2, 2016 from http://publicpolicy.airbnb.com/new-report-impact-airbnb-middle-class-income-stagnation/

Streitfeld, D. (2014, October 15). Airbnb Listings Mostly Illegal, New York State Contends. *New York Times*. Retrieved March 2, 2016 from http://www.nytimes.com/2014/10/16/business/airbnb-listings-mostly-illegal-state-contends.html?_r=0

Swisher, R. R., Kuhl, D. C., & Chavez, J. M. (2013). Racial and ethnic differences in neighborhood attainments in the transition to adulthood. *Social Forces*, *91*(4), 1399–1428.

Thebault-Spieker, J., Terveen, L. G., & Hecht, B. (2015). Avoiding the South Side and the Suburbs: The Geography of Mobile Crowdsourcing Markets. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing* (pp. 265–275).

Zimmer, M. (2010). "But the data is already public": on the ethics of research in Facebook. *Ethics and Information Technology*, *12*(4), 313–325.

Figure 1: The Types of Properties Listed on Airbnb, Broken Down by the Kind of Hosts

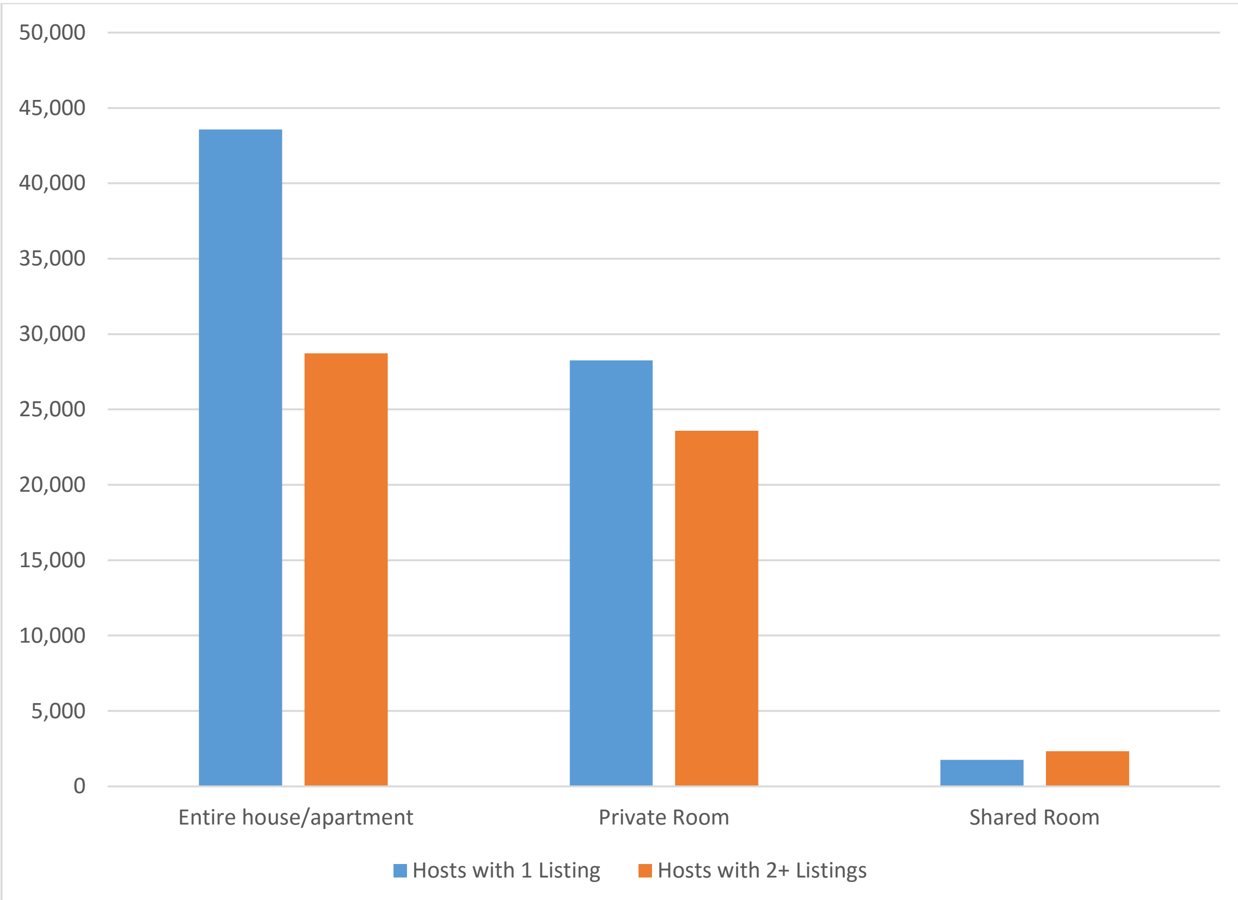

Figure 2: Average Nightly Price of Airbnb Listings by Property Type

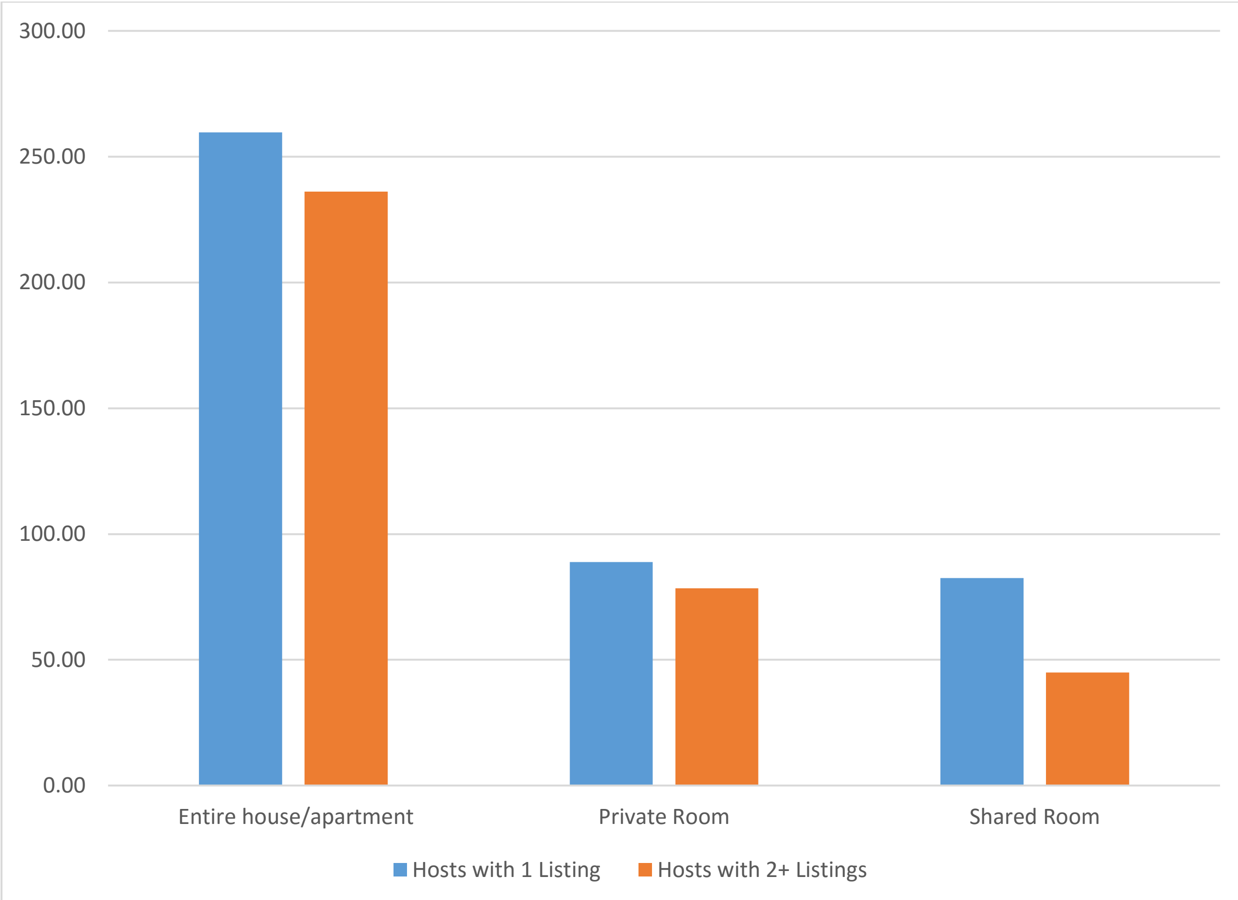

Figure 3: Average Nightly Price Per Person of Airbnb Listings by Property Type

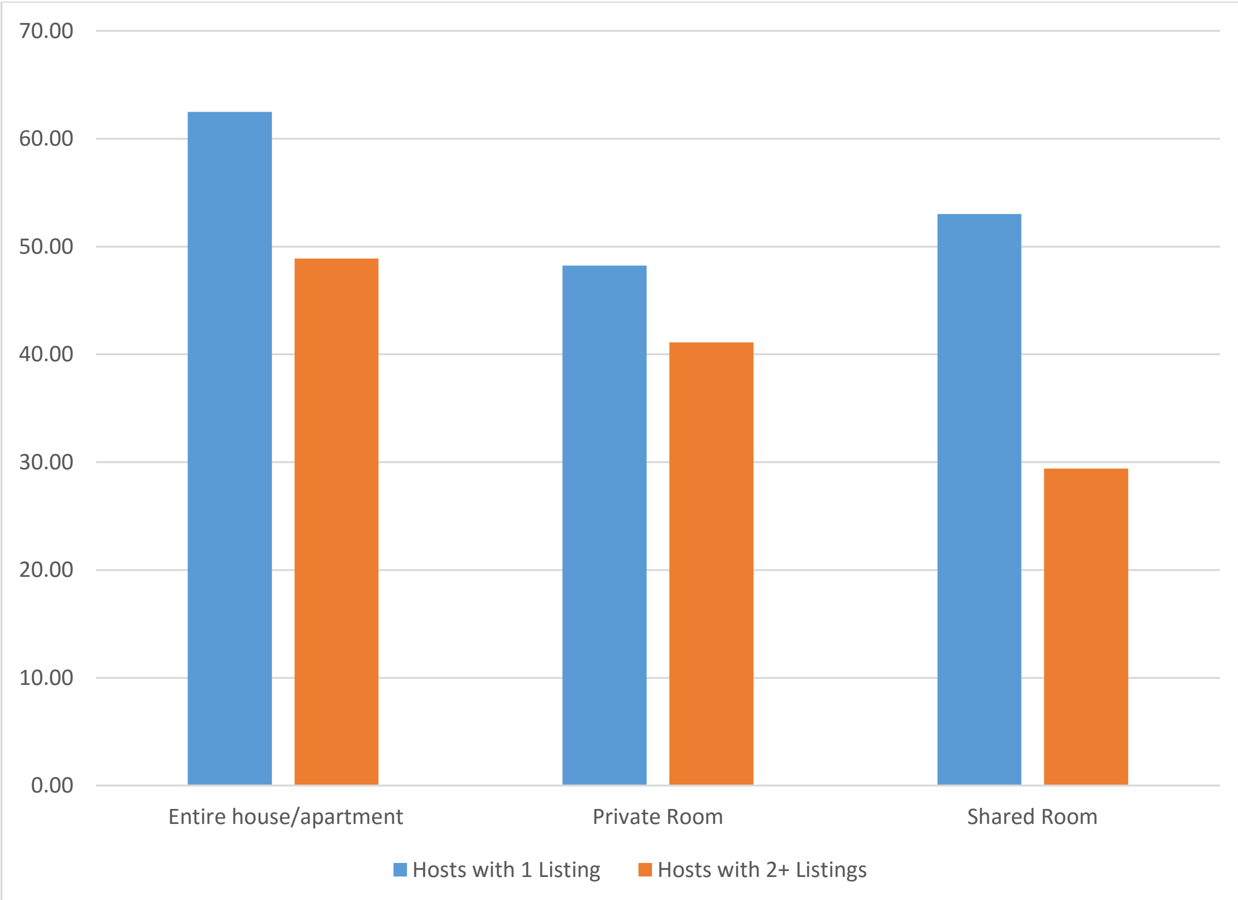

Figure 4: Booking Types by Host Type

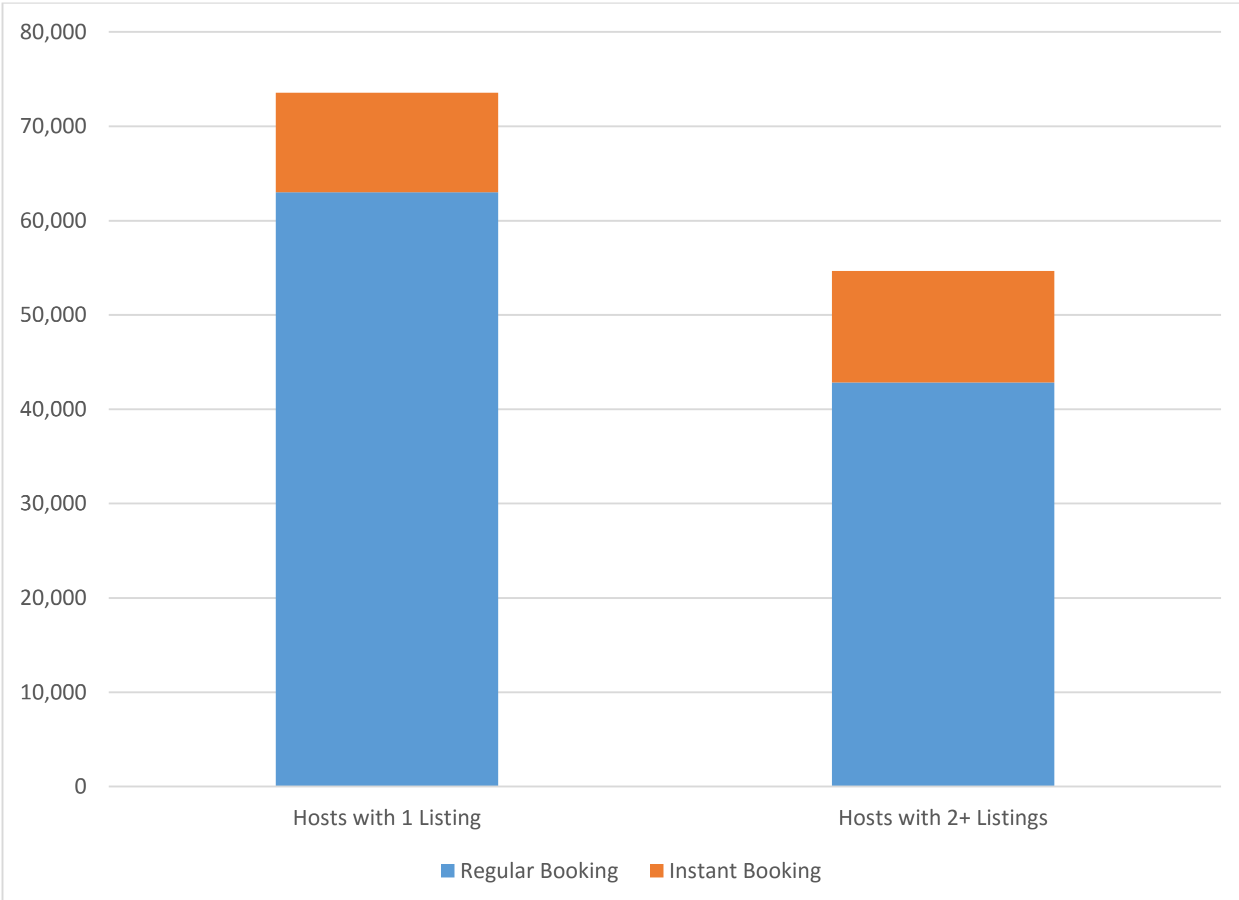

Table 1 – Number of Listings in Census Tract, Modeled with Mixed-Effects Negative Binomial Regression+

	Model 1	*Model 2*	*Model 3*
Census-Tract Level			
Constant	3.5442***	1.8606***	1.9114***
Population	1.0229***	1.1205***	1.1102***
Median Income	-	0.4309***	0.4236***
% with BA	-	4.6137***	4.7171***
% non-White	-	1.5942***	-
% Black	-	-	1.2240***
% Hispanic	-	-	1.5074***
MSA Level			
Population	0.9202***	0.8630***	0.8707***
MSA-Level Variance	4.9033	3.5329	3.5924
Deviance	160201.00	148156.86	148380.75
N	47655	47138	47138
# of groups	104	104	104
Average Group Size	458.2	453.3	453.3
AIC	160211	148172.9	148398.8
BIC	160254.9	148242.9	148477.6

* p<0.05, **p<0.01, ***p<0.001

+ The numbers reported for independent variables are incidence rate ratios

Table 2: Nightly Price of Listings, Modeled with Mixed-Effects Linear Regression

	Model 1	*Model 2*	*Model 3*	*Model 4*
Listing Level				
Constant	4.5469***	4.5488***	4.5798***	4.5797***
Length of Membership	-	0.0081***	0.0069***	0.0069**
Number of Amenities in Listing	-	0.1026***	0.0985***	0.0986***
Census-Tract Level				
Population	-	-0.0545***	-0.0571***	-0.0588***
MSA population	-	-0.0185***	0.0024	-0.0031
Median Income	-	-	0.0573***	0.0555***
% with BA	-	-	0.0551***	0.0492***
% non-White	-	-	-0.0756***	-
% Black	-	-	-	-0.0518***
% Hispanic	-	-	-	-0.0596***
Residual Variance	0.4585	0.4505	0.4515	0.4516
Tract-Level Variance	0.2064	0.1989	0.1716	0.1721
Tract-Level Variance as % of Total Variance	0.3104	0.3063	0.2754	0.2759
Deviance	283282.76	280816.06	278596.66	278634.46
N	128214	128214	127782	127782
# of groups	19227	19227	19127	19127
Average Group Size	6.7	6.7	6.7	6.7
AIC	283288.8	280830.1	278616.7	278656.5
BIC	283318	280898.4	278714.2	278763.8

* p<0.05, **p<0.01, ***p<0.001

Table 3: Nightly Price per Person, Modeled with Mixed-Effects Linear Regression

	Model 1	*Model 2*	*Model 3*	*Model 4*
Listing Level				
Constant	3.5148***	3.5172***	3.5716***	3.5700***
Length of Membership	-	-0.0075***	-0.0098***	-0.0096***
Number of Amenities in Listing	-	-0.0257***	-0.0302***	-0.0303***
Census-Tract Level				
Population	-	-0.0401***	-0.0392***	-0.0392***
MSA Population	-	0.0491***	0.0409***	0.0477***
Median Income	-	-	0.0344***	0.0285***
% with BA	-	-	0.1406***	0.1299***
% non-White	-	-	0.0056	-
% Black	-	-	-	-0.0155***
% of Hispanic	-	-	-	-0.0108**
Residual Variance	0.3068	0.3061	0.3071	0.3071
Tract-Level Variance	0.1349	0.1312	0.1028	0.1027
Tract-Level Variance as % of Total Variance	0.3054	0.3000	0.2508	0.2506
Deviance	231502.36	230927.18	228106.92	228087.92
N	128214	128214	127782	127782
# of groups	19227	19227	19127	19127
Average Group Size	6.7	6.7	6.7	6.7
AIC	231508.4	230941.2	228126.9	228109.9
BIC	231537.6	231007.5	228224.5	228217.3

* p<0.05, **p<0.01, ***p<0.001

Table 4: Number of Reviews per Listing, Modeled with Mixed-Effects Negative Binomial Regression[+]

	Model 1	*Model 2*	*Model 3*
Listing Level			
Constant	0.0283***	0.0288***	0.0288***
Number of Amenities in Listing	1.0605***	1.0620***	1.0608***
Census-Tract Level			
Population	0.9201***	0.9366***	0.9295***
MSA Population	1.1169***	1.1066***	1.0945***
Median Income	-	0.8961***	0.8994***
% with BA	-	1.1484***	1.1771***
% non-White	-	1.0530***	-
% Black	-	-	1.0244**
% of Hispanic	-	-	1.1006***
Tract-Level Variance	0.2971	0.2974	0.2972
Deviance	799566.76	796494.64	796408.20
N	128127	127695	127695
# of groups	19221	19121	19121
Average Group Size	6.7	6.7	6.7
AIC	799578.8	796514.6	796430.2
BIC	799637.3	796612.2	796537.5

* p<0.05, **p<0.01, ***p<0.001

[+] The numbers reported for independent variables are incidence rate ratios

Table 5: Listing Rating, Modeled with Mixed-Effects Negative Binomial Regression[+]

	Model 1	*Model 2*	*Model 3*
Listing Level			
Constant	5.2583***	5.2252***	5.2248***
Number of Amenities in Listing	0.8347***	0.8402***	0.8401***
Length of Membership	1.0176***	1.0200***	1.0204***
Census-Tract Level			
Population	0.9853**	0.9910	0.9979
MSA Population	1.0992***	1.0746***	1.0939***
Median Income		0.9501***	0.9483***
% with BA	-	1.0209*	1.0053
% non-White	-	1.1002***	-
% Black	-	-	1.0578***
% of Hispanic	-	-	1.0346***
Tract-Level Variance	0.0688	0.0642	0.0655
Deviance	476656.2	474500.28	474591.86
N	87401.00	87103	87103
# of groups	15943	15865	15865
Average Group Size	5.5	5.5	5.5
AIC	476670.2	474520.3	474613.9
BIC	476735.8	474614	474717

* p<0.05, **p<0.01, ***p<0.001

[+] The numbers reported for independent variables are incidence rate ratios

Table 6: Imputed Income, Modeled with Mixed-Effects Linear Regression

	Model 1	*Model 2*	*Model 3*	*Model 4*
Listing Level				
Constant	6.9284***	6.9016***	6.9477***	6.9500***
Length of Membership	-	0.5303***	0.5271***	0.5266***
Number of Amenities in Listing	-	0.1703***	0.1643***	0.1638***
Census-Tract Level				
Population	-	-0.1267***	-0.1089***	-0.1172***
MSA population	-	0.0818***	0.1067***	0.0869***
Median Income	-	-	-0.0401***	-0.0397***
% with BA	-	-	0.1508***	0.1674***
% non-White	-	-	-0.0818***	-
% Black	-	-	-	-0.0561***
% of Hispanic	-	-	-	-0.0162
Residual Variance	2.1588	1.9110	1.9139	1.9139
Tract-Level Variance	0.4759	0.3655	0.3233	0.3251
Tract-Level Variance as % of Total Variance	0.1806	0.1606	0.1445	0.1452
Deviance	329295.4	317625.08	315969.34	316001.6
N	88598.00	88598	88296	88296
# of groups	16069	16069	15990	15990
Average Group Size	5.5	5.5	5.5	5.5
AIC	329301.5	317639.1	315989.3	316023.7
BIC	329329.6	317704.8	316083.2	316127

* p<0.05, **p<0.01, ***p<0.001

Appendix 1: List of MSAs

New York-Newark-Jersey City, NY-NJ-PA Metro Area

Los Angeles-Long Beach-Anaheim, CA Metro Area

Chicago-Naperville-Elgin, IL-IN-WI Metro Area

Dallas-Fort Worth-Arlington, TX Metro Area

Houston-The Woodlands-Sugar Land, TX Metro Area

Philadelphia-Camden-Wilmington, PA-NJ-DE-MD Metro Area

Washington-Arlington-Alexandria, DC-VA-MD-WV Metro Area

Miami-Fort Lauderdale-West Palm Beach, FL Metro Area

Atlanta-Sandy Springs-Roswell, GA Metro Area

Boston-Cambridge-Newton, MA-NH Metro Area

San Francisco-Oakland-Hayward, CA Metro Area

Phoenix-Mesa-Scottsdale, AZ Metro Area

Riverside-San Bernardino-Ontario, CA Metro Area

Detroit-Warren-Dearborn, MI Metro Area

Seattle-Tacoma-Bellevue, WA Metro Area

Minneapolis-St. Paul-Bloomington, MN-WI Metro Area

San Diego-Carlsbad, CA Metro Area

Tampa-St. Petersburg-Clearwater, FL Metro Area

St. Louis, MO-IL Metro Area

Baltimore-Columbia-Towson, MD Metro Area

Denver-Aurora-Lakewood, CO Metro Area

Charlotte-Concord-Gastonia, NC-SC Metro Area

Pittsburgh, PA Metro Area

Portland-Vancouver-Hillsboro, OR-WA Metro Area

San Antonio-New Braunfels, TX Metro Area

Orlando-Kissimmee-Sanford, FL Metro Area

Sacramento--Roseville--Arden-Arcade, CA Metro Area

Cincinnati, OH-KY-IN Metro Area

Kansas City, MO-KS Metro Area

Las Vegas-Henderson-Paradise, NV Metro Area

Cleveland-Elyria, OH Metro Area

Columbus, OH Metro Area

Indianapolis-Carmel-Anderson, IN Metro Area

San Jose-Sunnyvale-Santa Clara, CA Metro Area

Austin-Round Rock, TX Metro Area

Nashville-Davidson--Murfreesboro--Franklin, TN Metro Area

Virginia Beach-Norfolk-Newport News, VA-NC Metro Area

Providence-Warwick, RI-MA Metro Area

Milwaukee-Waukesha-West Allis, WI Metro Area

Jacksonville, FL Metro Area

Memphis, TN-MS-AR Metro Area

Oklahoma City, OK Metro Area

Louisville/Jefferson County, KY-IN Metro Area

Richmond, VA Metro Area

New Orleans-Metairie, LA Metro Area

Raleigh, NC Metro Area

Hartford-West Hartford-East Hartford, CT Metro Area

Salt Lake City, UT Metro Area

Birmingham-Hoover, AL Metro Area

Buffalo-Cheektowaga-Niagara Falls, NY Metro Area

Rochester, NY Metro Area

Grand Rapids-Wyoming, MI Metro Area

Tucson, AZ Metro Area

Urban Honolulu, HI Metro Area

Tulsa, OK Metro Area

Fresno, CA Metro Area

Bridgeport-Stamford-Norwalk, CT Metro Area

Worcester, MA-CT Metro Area

Albuquerque, NM Metro Area

Omaha-Council Bluffs, NE-IA Metro Area

Albany-Schenectady-Troy, NY Metro Area

Bakersfield, CA Metro Area

Greenville-Anderson-Mauldin, SC Metro Area

New Haven-Milford, CT Metro Area

Knoxville, TN Metro Area

Oxnard-Thousand Oaks-Ventura, CA Metro Area

El Paso, TX Metro Area

McAllen-Edinburg-Mission, TX Metro Area

Allentown-Bethlehem-Easton, PA-NJ Metro Area

Baton Rouge, LA Metro Area

Dayton, OH Metro Area

Columbia, SC Metro Area

North Port-Sarasota-Bradenton, FL Metro Area

Greensboro-High Point, NC Metro Area

Little Rock-North Little Rock-Conway, AR Metro Area

Charleston-North Charleston, SC Metro Area

Stockton-Lodi, CA Metro Area

Akron, OH Metro Area

Colorado Springs, CO Metro Area

Cape Coral-Fort Myers, FL Metro Area

Boise City, ID Metro Area

Syracuse, NY Metro Area

Winston-Salem, NC Metro Area

Wichita, KS Metro Area

Lakeland-Winter Haven, FL Metro Area

Madison, WI Metro Area

Ogden-Clearfield, UT Metro Area

Springfield, MA Metro Area

Des Moines-West Des Moines, IA Metro Area

Deltona-Daytona Beach-Ormond Beach, FL Metro Area

Toledo, OH Metro Area

Augusta-Richmond County, GA-SC Metro Area

Jackson, MS Metro Area

Provo-Orem, UT Metro Area

Harrisburg-Carlisle, PA Metro Area

Scranton--Wilkes-Barre--Hazleton, PA Metro Area

Palm Bay-Melbourne-Titusville, FL Metro Area

Youngstown-Warren-Boardman, OH-PA Metro Area

Chattanooga, TN-GA Metro Area

Durham-Chapel Hill, NC Metro Area

Spokane-Spokane Valley, WA Metro Area

Lancaster, PA Metro Area

Modesto, CA Metro Area

Portland-South Portland, ME Metro Area

Appendix 2: Descriptive Statistics

	Nightly Price	Nightly Price/Person	# of Reviews
Mean	179.2456	52.1655	14.0466
Std. Deviation	321.5969	101.2377	28.5606
Min	10	0.625	0
Max	8000	8000	472
N	121968	121968	121886

	Rating	Imputed Income	Length of Membership
Mean	94.3556	4361.3730	23.6205
Std. Deviation	7.6794	8534.3950	18.0314
Min	20	10	1
Max	100	100000	98
N	83644	84764	121968

	# of Amenities	Median Income	% with BA
Mean	9.1373	66657.6100	49.6151
Std. Deviation	3.2004	32817.4400	21.9629
Min	0	2500	0
Max	20	250000	100
N	121968	121584	121679

	% non-White	% Black	% Hispanic
Mean	0.4350	0.1353	0.1781
Std. Deviation	0.2566	0.1946	0.1921
Min	0	0	0
Max	1	1	0.9851
N	121638	121638	121638

	Total # of Listings
Mean	2.6905
Std. Deviation	8.738
Min	0
Max	312
N	47655