Improving the Transparency of the Sharing Economy

Mathias Lecuyer*
Columbia University
New York, NY, USA
mathias@cs.columbia.edu

Max Tucker*
Columbia University
New York, NY, USA
mlt2158@columbia.edu

Augustin Chaintreau
Columbia University
New York, NY, USA
augustin@cs.columbia.edu

ABSTRACT

The idealistic beginnings of the sharing economy made ways to an entrenched battle to win over the public opinion and for law makers to appreciate its benefits and its risks. The stakes are high as the success of services like Airbnb reveals that under-utilized assets (e.g. spare rooms or apartments left vacant) can be efficiently matched to individual demands to generate a significant surplus to their owners. Rules and regulation, which are increasingly felt as necessary by many communities, also create friction over the best way to leverage these opportunities for growth. To make things worse, the sharing economy is complex and poorly documented: Three recent reports from public institutions and lobbying groups arrived at opposite conclusions with seemingly contradictory facts about the occupancy distribution.

In this paper, we show how to overcome this opacity by offering the first large-scale, reproducible study of Airbnb's supply *and* transactions. We devised and deployed frequently repeated crawls using no proprietary data. We show that these can be used to accurately estimate not only the supply of available rooms, but the effective transactions, occupancy, and revenue of hosts. Our results provide the first complete view of the occupancy and the distribution of revenue, revealing important trends that generalize previous observations. In particular we found that previous observations that seemed at odds are all explained by a variant of the "inspection paradox". We also found from our detailed data that enforcing a maximum occupancy of 90 nights a year would greatly reduce most concerns raised by various advocacy groups, while affecting only marginally the justifying claims that Airbnb quotes to argue for its beneficial impact.

Keywords

Sharing Economy; Airbnb; Measurement; Transparency

*Both primary authors indicated in alphabetical order.

WWW 2017 Companion, April 3–7, 2017, Perth, Australia.
ACM 978-1-4503-4914-7/17/04.
http://dx.doi.org/10.1145/3041021.3055136

1. INTRODUCTION

Matching platforms like Airbnb, Uber, and Kickstarter went in a few years from niche idealistic beginnings to mainstream use - some would even argue abuse - with increasing impact. The core principle of the "sharing economy" is simple: owners of various *underused* assets (*e.g.* a room left unoccupied, an extra seat in a car and some spare time, or some capital) ought to remain free to make a deal to satisfy the large occasional demand for those goods (*e.g.* to accommodate or transport travelers, to support entrepreneurs). The fact that such transactions take place (at a price satisfying both the occasional demand and the owner) confirms *in retrospect* that there is an untapped "surplus" that would be wasted in a more traditional arrangement where the owner is the sole person enjoying access to this asset. That the aforementioned surplus emerges from flexibility is a key argument used by advocates of the sharing economy to criticize new rules and proposed regulations. Those advocates are also quick to point out that a significant fraction of owners of those underused assets are from middle class families complementing their income [1], or from neighborhoods that are currently left out from hospitality services and the associated economic development [15].

Here we focus on the Airbnb matching platform connecting home owners and renters with occasional travelers for accommodation. The controversy over Airbnb's widespread use in metropolitan centers like San Francisco, London, and New York, best illustrates the limitations of the current status quo, often operating outside the law[1]. Listings from New York City alone were reported to generate up to half a billion dollars in yearly transactions, while national trends suggest that this total increases by a half every year [14]. Airbnb is a growing global force now valued at $30 billion, but it currently faces multiple challenges: Enforcing regulations so its hosts do not discriminate in a way that creates a disparate impact for vulnerable populations [7], addressing concerns that its service replaces traditional hotel offerings in some communities [20], or is argued to be a revenue loss for municipal governments estimated in the tens of millions [14].

One chief concern is a lack of transparency and accountability. Traditional hospitality businesses file and report activities, which allows for urban zoning and taxation. In contrast, Airbnb operates opaquely, just as its intermittent spread has had an impact on the lifestyle and the availability of affordable housing in many neighborhoods. Because of this opacity, recent developments in the news signal dis-

[1]It has been reported that a majority of short term rentals are *de facto* illegal under New York State law [17, 16].

trust surrounding the data, and whether reported facts are truthful or representative. Prior to our work, both data released by Airbnb about the impact of its service in the New York community [1], and a study based on other proprietary data [14] received harsh attacks about their mere factual honesty. In an independent report [6], Airbnb's data snapshot was criticized as "photo-shopped" and it was argued that the company resorted to a one-time purge to "ensure that it would paint a flattering picture". Conversely another study, funded by the American Hotel & Lodging Association, was publicly described by an Airbnb spokesman [9] as "The hotel industry gets what it pays for, which in this case is a specious study intended to mislead and manipulate.".

In this paper we present a carefully designed methodology, based on repeated measurements, that allows us to infer both available listings and transactions taking place on Airbnb's platform (Section 2). We also describe our system implementing this methodology, which enables large scale, longitudinal measurements of Airbnb's sharing economy. We use this measurement data to validate our methodology by comparing our results to publicly released aggregate numbers of offer and demand on Airbnb (Section 3). In addition to making publicly available the first data-set on bookings and host revenues on this platform, we reveal important trends that were left unnoticed by previous works.

- We carefully evaluate the highly debated issue of hosts owning multiple listings, and its evolution over time. While previous proprietary studies omit details and temporal trends, we confirm that the small minority of multi-listing owners collectively receive a significant share of the overall revenue in New York City. The year 2015 appears as a turning point; it calls for caution in interpreting previous claims. (Section 4)
- We report for the first time the full distribution of occupancy ratio and its effect on revenue. This analysis reveals a simple statistical bias akin to the well known inspection paradox that explains why previous claims, used in different contexts, seem contradictory. Depending on the point of observation (as a platform or as a traveler) the prominence of occasional bookings vary from a typical case to a rare occurrence. In terms of income, however, the lion's share of revenue goes to a minority of listings which are offered and booked very frequently. (Section 5)
- These observations lead us to revisit Airbnb's claims about how to regulate its use without adversely impacting the complementary income of middle class families. The new details of our data suggest that enforcing the proposed 90-day or 30-day maximum yearly occupation is effective in preventing the revenue from the sharing economy to become entirely captured by a minority of dedicated listings, while leaving a majority of hosts unaffected. It should however be noted that the impact on the total revenue, and the overall supply of listings, is dramatic. (Section 6)

2. AIRBNB REPEATED MEASUREMENTS

Collecting information from Airbnb poses new challenges, which most previous efforts ignore by focusing on the supply side of the market. We briefly highlight the limitations of those methods, including recent ones not presented in peer-reviewed work. We then present our repeated scraping method, to our knowledge the first one applied to Airbnb, that allows us to collect calendar data on all listings in a given city at short intervals, as well as our methodology to estimate the demand side of Airbnb's marketplace. We carefully corroborate this methodology's results in §3.

2.1 The Challenges of Transaction Dynamics

Airbnb is a two-sided marketplace that matches demand for short term housing to a set of hosts making listings of various types available at multiple dates. To meet a highly dynamic demand from travelers, Airbnb must quickly respond to various detailed online queries. Measuring at scale the *supply* of the marketplace is hence made easy by simultaneously performing parallel queries separated in space. Because new listings do not appear very fast, it is possible to follow trends with a few data collections, and aggregate supply by merging various availabilities. That is the method of choice behind all independent measurement efforts reported so far [19, 15, 6, 21, 20].

Measuring the *demand* and the dynamics of *transactions* taking place on this marketplace is much more difficult, since booking information is not directly available. All previous reports have been done using proprietary data, either obtained internally [1], through legal means via subpoenas [16] or from another company with undisclosed methods [14]. The restrictions in scope and legitimacy of those methods render the sharing economy essentially opaque, a situation that our method intends to drastically change. The only alternative we found outside of our work exploits the number of reviews in order to infer the number of nights booked for a particular listing [5]. As we show in section 3.2, this method adds a significant variation that correlates with other attributes, introducing a systematic bias and other drawbacks.

Here we measure transactions and demand by repeating *in time* measurements of calendar availability for each listing observed, and by *combining* information from multiple dates together. Thus we can leverage much more information about availability dynamics and hence infer transactions taking place. The immediate challenge is the scale of data collection (described immediately below), but we show that it can be overcome. As evidence, we implemented our method on one of the densest local markets (listings located in Manhattan). The second challenge is that in spite of leveraging much more data, we are still subject to inaccuracies coming from missing information. We found in practice that a few heuristics and filters (§2.3) on the interpretation of dynamic availability can be used to detect and ignore unreliable calendar data. Section 3 shows that these heuristics reduce variation over time for different aggregates, and that our estimates of these aggregates always agree with previously reported values made with proprietary data.

2.2 Methods for Repeated Data Collection

Figure 1 shows the architecture of our scraper. Each of our servers has multiple IP addresses, exposed to workers through a proxy. The proxy binds all interfaces corresponding to each IP, and forwards incoming requests through the incoming interface. Workers thus only need to contact the proxy through the desired IP address, and the request to Airbnb will use the same address. We rate limit the workers at the granularity of each IP, to minimize the number of HTTP errors and reduce the load on Airbnb.

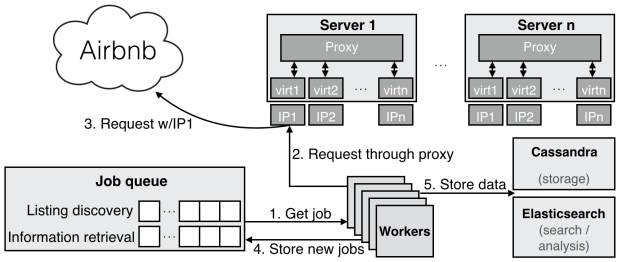

Figure 1: Scraper's architecture. Workers are programs makings requests to Airbnb to discover listings and collect calendar data. They coordinate through a job queue.

The input to this infrastructure is a list of cities to be scraped. Each scrape consists in two kinds of jobs that query public information from Airbnb, through the API endpoints used by their web and mobile applications. These endpoints are accessed without needing an account. We next describe each job type, and their usefulness in the data collection process.

Listing Discovery.

To collect data on available listings' calendars, we first need to enumerate all, or at least most, listings available in a given city. That is challenging since Airbnb does not disclose an exhaustive list of active listings, some listings come and go out of market and, like in most online services, a cap is maintained on the number of results and number of queries that a machine can perform.

However, we found in practice that one can exploit multiple aspects of the platform to end up with a quasi-exhaustive list. First, there is a rich set of features: *e.g.* neighborhoods, types of listings, size, and price. Hence when a given query in a neighborhood has more than 1000 listings, which is common in New York, we refine the query with listing types (full apartment, and private or shared room), and a price range that we split in two until there is no more than 1000 listings found. Second, the platform offers various exploration tools: *e.g.* similar listings, reviews from users, geographic search, which we use to expand the search beyond the strict definition of a particular query. This helps in discovering listings as well as new features such as new names of neighborhoods to crawl. Third, we can leverage memory from past measurements to amortize the exploration phase. However we found one must also do this carefully: In practice maintaining a log of all listings previously seen and re-querying all of those each time is counter-productive. Since a fraction of those became inactive one ends up with a high rate of failed queries, which are also suspicious. A good middle ground is to leverage memory from past explorations at the feature level (*i.e.* reuse names from all neighborhoods previously discovered) which always provides sound results.

Combining these techniques, our discovery process quickly finds all neighborhoods associated with a given city. Each subsequent crawl takes less than one day, enabling frequent measurements, and we observe more than 90% of all listings found in prior works around the same time period. Of course, it is still possible that we miss a listing during a longer period despite the repeated measurements. Our heuristics carefully account for this when analysing data. Note moreover that this implies that the missing listing did not appear for multiple days in a detailed search, which makes it doubtful than anyone else would see it on that day to perform a booking in the next year.

Information Retrieval.

For a given listing id, we query the "listing" and "calendar" endpoints. The former needs to be done only once and returns information such as the property type, property category, approximate location, and various amenities offered. It also includes a unique id for the host renting the listing, permitting us to study hosts who manage multiple listings under the same account.

The calendar endpoint returns, for the next 12 months (including the current month), the daily availability status and price in US dollars. On each day of data collection d_c, for each listing l in a given city, and for each *target day* $d_c < d_t \in T$, we observe the future availability $S_l(d_c, d_t) \in \{0, 1\}$. When the scrape performed on d_c indicates that a property is available for booking on d_t, we say $S_l(d_c, d_t) = 1$. Otherwise we write $S_l(d_c, d_t) = 0$. In addition to observing the availability status of the listing, we check the price of any available listings for each open calendar day, and update the most recent price observed for any target day.

2.3 Inferring Bookings from Repeated Views

One of the main contributions of this paper is our methodology to infer bookings and revenue from repeated measurements. We next describe this methodology, and we will show in §3 that it allows us to accurately reconstruct aggregates previously released by Airbnb, including the median occupancy ratio, the percentage of listings shared for less than 90 or 120 days per year, and the hosts' median income.

For each Airbnb listing, each target calendar day d_t is seen at multiple times in the past. The availability of a listing for that night evolves over time, and we exploit this information to infer booked nights. For any target day d_t we formally assign a listing to one of four states:

A *Available.* Denotes a listing that was offered on the market but was left unoccupied.

B *Booked.* On that day the listing was offered on the market and rented out to a guest.

U *Unavailable.* The listing was not proposed on the market and hence also not booked by anyone.

N *No Data.* Denotes a listing for which information is not sufficient to infer its booking and availability status with confidence.

We introduce for every listing l and target day d_t its latest observation period, *i.e.* $L_l(d_t) = \max\{d_c | (d_c, d_t) \in \text{Domain}(S_l)\}$, if there is no such d_c, we set $L(d_t) = -\infty$. For that listing on that target day, we automatically classify the listings in state N if $L(d_t)$ is undefined or $|L(d_t) - d_t| > \Delta$. In other words, this *filtering heuristic* that we apply excludes data from target days that were not observed in the recent past, to prevent multiple issues: First, transient listings that appear infrequently or stop appearing during the observation window might be classified as available for a long period in the future, without later observations confirming this availability. That could wrongly classify occasional renters into full time operators and which ought to be avoided. Second, if the listing was not removed from the market, but our method fails to observe it, it is more likely that a booking took place in the interval we missed, overestimating the amount of nights unoccupied.

Formally, we denote by $X_l(d_t)$ the state of this listing for that target date. Assuming that $|L(d_t)-d_t| \leq \Delta$ (otherwise, by the convention above $X_l(d_t) = N$), we set $X_l(d_t)$ to

$$\begin{cases} A \text{ iff } S(L(d_t), d_t) = 1, \\ U \text{ iff } \forall d_c \leq d_t,\ S(d_c, d_t) = 0, \\ B \text{ iff } S(L(d_t), d_t) = 0 \text{ and } \exists d_c < d_t,\ S(d_c, d_t) = 1. \end{cases}$$

In other words, a listing that was never observed as available for that night is set in state U. A listing that is available for a night d_t during its last observation $L(d_t)$ is in state A (except for handful of cases, those were also continuously available at all previous observations in the period). A listing is deemed booked hence in state B only if that night was previously observed as available in the past and is not available *any more* during the final observation.

Note that these rules might misclassify the true state of a booking. First, we may *underestimate* the amount of nights booked, since any target day in our observation period that was booked prior to the beginning of our measurement is going to be wrongly assumed to be unavailable (hence in state U). That effect must be particularly strong towards the beginning of our observation period. As expected, we do observe that unavailable listings are more common in the beginning but we see this effect disappear after roughly two weeks. We did not include those early target days to account for this effect. Second, we can possibly *overestimate* the amount of bookings if a host decides, during our observation period, to remove a listing from the market for independent reasons (e.g. a change in vacation plans, another event). The importance of that effect is indeed hard to quantify, but computing occupancy using the above heuristic, we found that our estimates match remarkably well the publicly available percentile values released by Airbnb for the same observation period (see bellow). Note moreover that this issue does not apply when hosts plan their availability in advance, or inversely when they do not update their calendar with last minute unavailabilities (since they can always refuse a booking).

To translate booking nights into gross revenue, we simply assume that a night generates revenue only if it is in the B state, at a price that is the most recent price offered for that day when it was still available. This allows us to account for price variations over different periods of the year, different days of the week, or over time. For instance, many listings are more expensive during weekends or vacation periods, and Airbnb even offers a tool to optimize prices according to the season, demand, and other factors [11].

Finally, with frequent enough measurements, we can estimate the start and end dates of entire stays. We identify stays by grouping consecutive days of a given listing that were detected as booked for the first time during the same crawl. We also make sure that the time span of these consecutive days cover a period longer than the minimum stay length of the listing. While we may inadvertently merge two different, but consecutive, stays that were both booked between the same two crawls, frequent measurements lessen this problem.

	listings			revenue
type	This paper	[1] 2015	[5] 2016	This paper
EH	65.2%	63.7%	58.8%	79.5%
PR	32.5%	36.2*%	38.2%	19.6%
SR	2.3%	N/A	3%	0.9%

(* this estimate includes shared rooms as well)

Table 1: Comparison of the supply of listings by listing type.

3. EXTERNAL VALIDATION

We implemented the architecture above at scale to reliably collect listings indicating "New York" as their home city, with observations gathered roughly every one or two days. We checked the zip-code for all of those listings and found that, with a handful of exceptions that were removed, they include listings located in Manhattan (other listings within New York city limits almost always indicate the borough as their home city). During each crawl we observed around 14k unique listings, more than 90% of the number observed for Manhattan in previous works [5, 14, 1, 16].

3.1 Comparison with Previous Reports

We consider two periods spanning different seasons, when our data collection was continuously operating: Apr. to Jul. 2015, and Oct. 2015 to Feb. 2016. The first one corresponds to the last 6 months that were used in Airbnb's own report [1]. We will use that report as our main point of comparison, while occasionally referring to other similar efforts when relevant, including the investigation by the Office of the Attorney General of the State of New York (NYAG) [16], a recent academic report with an undisclosed data collection method from 2015 [14], and an independent project (insideairbnb.com) that made monthly snapshots of Airbnb, but used a review based method to estimate bookings. By design, we can only compare metrics that those sources have reported on, as we cannot access the original data, and the only open source project (insideairbnb) does not have frequent enough crawls. However, the reported metrics are sufficiently rich and diverse to corroborate our approach.

Room Types.

Airbnb offers three general types of accommodation: "shared room" (SR), "private room" (PR) and "entire home/apartment" (EH). Table 1 shows a close match on the supply of listings in each category with other externally reported metrics. Our study confirms that shared rooms are a small minority. For the first time, we can also compare the revenue made by each category of listings. We observe that shared rooms account for an almost negligible share of the business, while entire homes take the lion's share of revenue.

Multi-listing Owners.

We refer to a host's *Multiplicity* as the number of listings they own. Hosts with multiple listings have drawn increased scrutiny from those who fear that home sharing might motivate individuals to purchase additional residential units with the intention of using them as de-facto hotels.

Table 2 shows the relative number of hosts owning multiple Entire Home listings as reported by Airbnb, by an independent project, and using our methods. These numbers match well, and appear robust over time. While Airbnb did not reveal the same distribution when listings from other

# owned listings	Airbnb [1] 2015 Inner Manh.	Outer Manh.	[5] 2016 Manh.	This paper `Sp-Su`	`Fa-Wi`
1	95.2	95.2	93.8	92.2	91.5
2	3.18	3.25	4.81	5.08	5.71
3	0.79	0.74	0.98	1.42	1.47
4	0.26	0.30	0.16	0.74	0.62
5	0.20	0.13	0.06	0.32	0.30
≥6	0.36	0.36	0.14	0.19	0.43

Table 2: Percentage of owners with multi-`EH` listings.

# owned listings	[16] -2014 NYC	[14] 2015 NYC	[5] 2016 Manh.	This paper `Sp-Su` Manh.	`Fa-Wi` Manh.
1	94.5†	85.7	89.4	87.7	85.6
2	N/A	9.2	7.7	8.29	9.80
≥ 3	5.5	5.1	2.9	4.02	4.57

(† also includes hosts with 2 listings)

Table 3: Percentage of owners with multi-`EH`,`PR` listings.

categories such as `PR` are included, this was done in the past, and we can again reproduce these numbers with high accuracy (see Table 3). We defer more discussion of this controversial issue to the immediate next section.

Occupancy.

Another metric of controversial use is the occupancy, which measures the fraction of nights that a unit is booked for short-term rental. Again, this metric is introduced as a hint that high occupancy denotes listings that are operated as illegal hotels. Based on our methodology, we compute the *occupancy ratio* for each listing by dividing the number of nights we observe in state B to the total number of nights in the period. The denominator includes many cases: if the listing was never available, if it was available and booked very late after our last observation point, or even when the listings ends up in the N state that denotes lack of conclusive evidence. We choose that convention to provide a conservative estimate of occupancy, one that cannot be inflated due to the limitations of our crawl. Note that some studies classify listings on the fraction of nights made available (as opposed to booked) [5, 14]. We choose not to use this metric as we observed that a listing can be available most of the year for many reasons, including the owner's not updating its calendar, without concluding that it is used year round.

There are very few publicly shared figures about occupancy. Airbnb claims [1] that in New York "The vast majority of listings are shared only occasionally. The median number of nights booked per listing in the past year is 42, with 84% of listings shared less than 120 days per year and 78% of listings shared for less than 90 days per year." Figure 2 presents the complete distribution of occupancy ratio among `EH` listings in Manhattan in our observations, where triangles indicate the numbers cited above. We found a very close match for the most relevant period `Sp-Su`, when we carefully limited our estimation to the most reliable data (by setting Δ to 5 days), with only a slight overestimation of occupancy for one data point. For the later period `Fa-Wi`, during which our crawl was able to operate more frequently and consistently in time, we found that the distribution is not affected by the value of Δ anymore. Occupancy seems

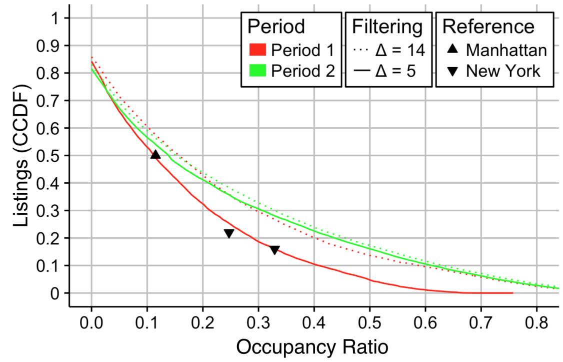

Figure 2: `EH` occupancy ratios in Manhattan. The triangles show occupancy numbers publicly released by Airbnb.

Δ	Airbnb [1] 2015 Outer	Inner	[6] 2016 Manh.	This paper `Sp-Su` 5	`Sp-Su` 14	`Fa-Wi` 14
	N/A		N/A			
`EH`	0.11	0.10	0.31*	0.11	0.15	0.15
`PR+SR`	0.13	0.15	0.30*	0.15	0.20	0.14
All	0.11	0.11	0.31*	0.12	0.17	0.14

(* denotes an average, other figures are medians)

Table 4: Comparison of reported occupancy ratios.

slightly higher in this case; this may relate to the Holiday season when many New Yorkers choose to travel.

Finally, we compare how occupancy varies with the type of listing (note that only medians were released by Airbnb) as seen in Table 4. We agree with Airbnb that `EH` listings are typically less occupied (and available) than rooms which are more frequently offered continuously.

Revenue.

Even less reliable public information exists for revenue. Since hosts' revenue and its distribution is the source of much contention between proponents and opponents of Airbnb, we dedicate a separate section to its study (§6). In this section, figure 7 shows that our methodology produces a median income close to the one released by Airbnb, the only public figure that we found.

3.2 The Biases of Review Based Methods

Some projects advocate using the number of reviews posted in order to infer the number of nights booked for a particular listing [5]. It works as follows: first the method assumes a given review rate, typically constant among listings, and an average number of nights for each stay. By multiplying the number of reviews seen publicly during a period with the inverse of the review rate and the average stay length, one can estimate the number of nights booked for each listing.

Using our rich dataset with repeated measurement, we can replicate and assess this methodology for comparison. First, we estimated the number of stays, and their length, as described in §2.3, from which we computed the average length of a stay. We also collected the reviews for each listing, allowing us to compute the booking rates and the average number of reviews per night booked.

Various estimates of the review rate have been used in the past: The value of 72% was used following a comment left by Airbnb's CEO on Quora [3] in 2012, while another report from the San Francisco Board of Supervisor (SFBOS) [4] used an estimate of 30.5% in 2015. The insideairbnb project

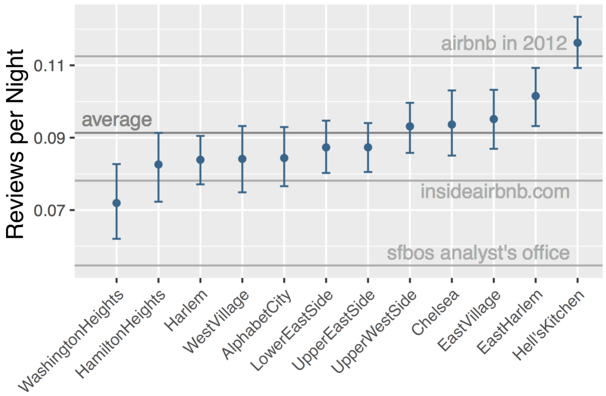

Figure 3: Review rates in different neighborhoods. The error bars show the 95% confidence interval.

settled for 50% as a good middle ground. We independently measured a review rate of 55% in Manhattan through our reproducible methodology. Airbnb reported an average stay length of 6.4 nights in New York [2], while our method observed it to be 6.1 nights. Thus, our method supports the average estimates used by prior art.

However, the more detailed analysis made possible with our data reveals that behind those average numbers, review based estimates are bound to create large inaccuracies. Figure 3 presents the number of reviews per night booked in Manhattan's neighborhoods with the most listings, together with 95% confidence intervals computed using the bootstrap method [8]. Previously used values are in the confidence intervals of only 5 neighborhoods. The average rate of reviews per night booked goes from 7.1% in Washington Heights to 11.6% in Hell's Kitchen. The mean of 9.1% is far from those, and only approximates well roughly 7 neighborhoods. A review based method to study the geographical impact of Airbnb, as [15] did, would underestimate the revenue in Washington Heights by 28% and overestimate it by 22% in Hell's Kitchen, even while using the exact borough average. Fortunately, using our repeated measurement method, this bias can be corrected.

	Price		Bookings		Listing type	
	Below median	Above median	Below median	Above median	Entire home	Private room
Reviews / stay	58.3%*	52.3%*	56%	55.1%	54.2%*	57.5%*
Avg stay length	5.8*	6.3*	5*	6.2*	6.3*	5.8*
Reviews / night	11.2%*	8.8%*	10%*	8.2%*	8.6%*	9.9%*

(* statistically significant mean difference, $p \leq 0.01$)

Table 5: Review rate and average stay length for different subgroups. Most subgroups have Statistically significantly different review rates and average stay lengths.

In addition to geographical bias, reviews show a *systematic bias* among listings, as Table 5 reveals. The correlation between review rates and room types has already been documented [12], and we confirm this trend. We also show that more expensive and more frequently booked listings present lower review rates and a longer average stay length. The differences are statistically significant, using a permutation test [10], and a method deducing revenue from reviews would thus overestimate the revenue of cheaper listings that are rarely booked by 27%.

Finally, we remark a couple of other drawbacks from review based methods: Since they rely on multiplying reviews by factors, estimated occupancy can exceed 100%. It must be corrected, for instance using a maximum occupancy rate, introducing yet another factor for noise. Review based methods also make it harder to account for varying prices, since we do not know what specific days were covered by the review. And finally, such a method is fragile, since a service like Airbnb can easily curate reviews to show only a subset of them, especially for listings containing many reviews, including old ones. Such a curation would minimally affect its business and certainly skew any possible estimate. In contrast, perturbing the results of our method is harder, as it would have to affect the calendar availability of listings to make them appear less available, which would result in a direct loss of business opportunities.

# owned listings	Airbnb [1] 2015	2016*	This paper `Sp-Su`	`Fa-Wi`
1	59	86	74.3	78.4
2	16	7	11.9	10.8
3	8	3	5.83	4.90
4	6	1	3.86	2.51
5	3	0	2.12	1.07
6+	7	2	2.06	2.26

(* projection based on November 17, 2015)

Table 6: Percentage of revenue to owners of multi-`EH` listings.

# owned listings	[16] -2014 NYC	[14] 2015 NYC	This paper `Sp-Su` Manh.	`Fa-Wi` Manh.
1	62.7†	68.0	67.4	68.5
2	N/A	15.5	15.9	16.0
≥ 3	37.3	16.5	16.7	15.5

(† also includes hosts with 2 listings)

Table 7: Percentage of revenue to multi-`EH`,`PR` listings.

4. MULTIPLICITY AND REVENUE

Multi-listing owners are unanimously seen as a minority. However, estimating their share of revenue was the subject of much bitter controversy: Analyzing bookings from 2010 to 2014, the New York state Attorney General [16] concluded that while only 6% of the hosts, those owning 3 or more listings accounted for 37% of the revenue. In November 2015 Airbnb released a report claiming that this trend was primarily a historical artifact: it reported that 24% of revenue went to those hosts in 2015 and projected based on current data that this share would reduce to 6% by the end of 2016. A recent report funded by Airbnb competitors found figures resembling those reported for 2015 [14], but did not disclose their data source.

Now, thanks to our reproducible method, multiple new claims can be made: (1) We independently established that both figures are accurate. In fact we can *reconcile* those observations: the discrepancy originates as Airbnb focuses on hosts owning multiple `EH` listings (see Table 6), while the other study includes other types of listings in multiplicity count (Table 7). Indeed, most of those listings are private

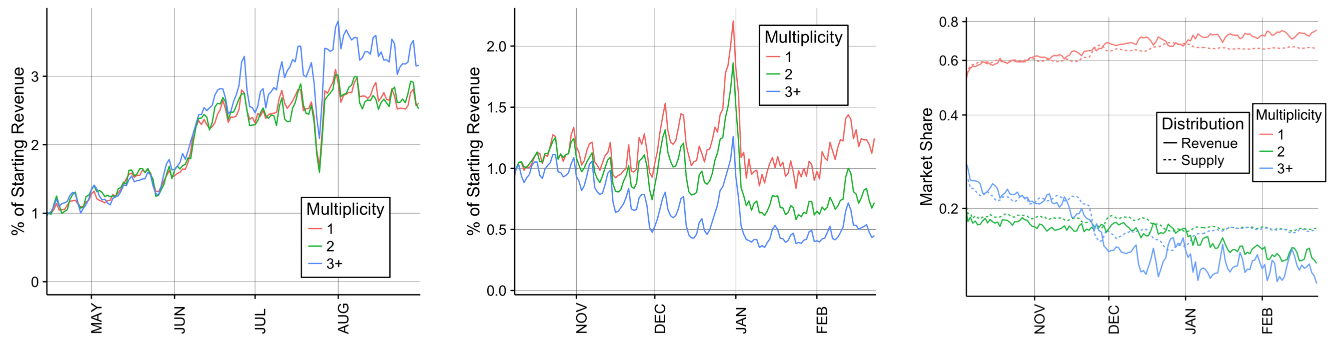

Figure 4: Revenue evolution for multi-`EH` listings: `Sp-Su` (left), `Fa-Wi` (middle), `Fa-Wi` relative share (right).

rooms, which falls under the same legal regime as entire home as they are considered "private" short term rentals under New York State Law [16]. (2) The announcement that revenue for multi-listing hosts is now drastically reduced is exaggerated in the short term - overall we found nearly no difference between the two periods - but it holds in the long term. As Figure 4 illustrates, absolute revenue expanded faster for multi-listing owners in `Sp-Su` but shrank during `Fa-Wi`. This is primarily a reduction in supply of listings with multiplicity that followed Airbnb's tightening of rules that was previously reported [18] and is now confirmed. Our reproducible method allows to double check that this trend continues, and may expand transparency by monitoring more types of listing. Finally (3) the claims that mega operators (with three or more listings) saw the largest revenue increase in the recent past and are hence an aggravating factor is an overstatement, at least for `EH` listings in New York.

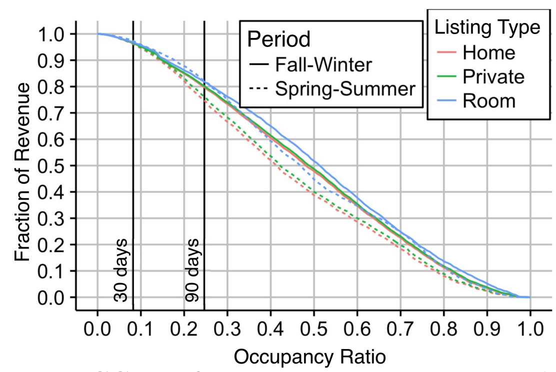

Figure 5: CCDF of revenue over occupancy ratio for different types of listings. We see that listings rented for more than 30 and 90 days per year capture 95% and 80% of the revenue respectively.

5. OCCUPANCY AND REVENUE

Occupancy and its effect on revenue is controversial, and figures advertised publicly seem *a priori* contradicting. On the one hand, Airbnb claims that a majority of its listings are shared only occasionally (stating that the median occupancy ratio is around 11%). On the other hand, competing studies claim that a minority of highly available listings (*e.g.* those made available at least 360 days a year - called full time operators) capture a substantial amount of revenue. We can reconcile those views as we provide the complete distribution of occupancy ratio. We show previous disagreements are simply the results of a statistical bias sometimes referred to as the "inspection paradox".

In Figure 5, we plot for every value of x, the fraction of revenue that goes to listings with occupancy ratio larger than x. We present this distribution for different listing types and periods, and find that it is very robust. Thanks to this observation we can draw important conclusions and reinterpret previous claims: (1) Relying on median occupation (around 11%) is misleading since all listings with up to this value account together for less than 5% of the revenue, because they gather roughly 5% of the bookings. To put it differently, a "typical" listing sees occupation around 11% and hence is shared 4 days a month on average. However, a "typical" *traveler* is likely to be associated with a listing booked around [40%-50%] of the time, since listings with occupation ratio above this value capture half of the total revenue. (2) The apparent discrepancy between various studies is hence a consequence of a variant of the *inspection or waiting time paradox*: the distribution of a variable observed during a booked night is not the same distribution as if a listing was picked at random. This statistical effect is prevalent for the type of skewed occupancy distribution that we measured. Many occasional listings exist, but nights booked for those places are all in a tail of rare events that are simply not likely to show up in an itinerary. (3) It is correct that the sharing economy concentrates revenue in a minority of active listings: previous studies claim that 24% of the revenue belongs to the 3% of full time operators. This is far from an isolated case: we found for instance that the 35% of listings that are booked at least 90 days capture 80% of the overall revenue.

Figure 6 presents the revenue seen across occupancy ratios for listings of different multiplicities. We see that multi-unit listings in general have a slightly higher occupancy. However, that effect is of secondary importance, and not a main driver of revenue.

6. REVENUE DISTRIBUTION

Finally, thanks to our detailed data, we can observe how revenue is distributed among hosts, a controversial topic for which no reliable distribution is known. We first describe in figure 7 the distribution of yearly income that we observed among all hosts, where we denote by two small black dots the only figures publicly shared by Airbnb: the median income in Inner Manhattan and Outer Manhattan. Let us first focus

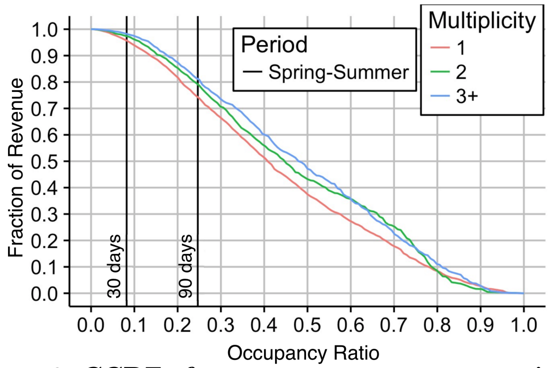

Figure 6: CCDF of revenue over occupancy ratio for different listing multiplicities.

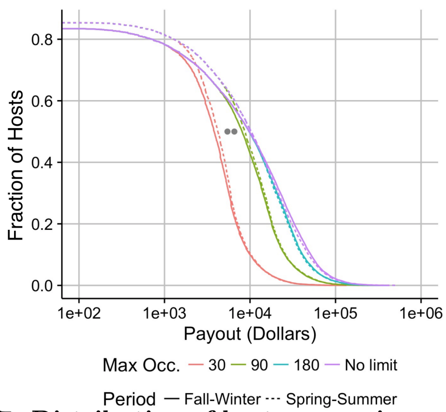

Figure 7: Distribution of hosts across income levels. The black dots correspond to the median income in Inner and Outer Manhattan shared by Airbnb.

on the purple line which denotes the yearly income projected from our observation window (the other curves will be described in the following paragraph). First, it is striking that our method independently produced median host incomes in the same range as those released by the company, and is quite robust in time. Second, we find that while many hosts make small incomes, those in the top 20%, 10%, or 5% receive orders of magnitude more in revenue. For that reason, it is informative to plot the distribution of revenue by income level. In Figure 8, we plot for any x the fraction of revenue going to hosts earning at least x. We observe that the same variants of the inspection paradox apply when classifying hosts per income: While a median Airbnb host in New York receives approximately \$10k a year, a traveler meets half of the time a host earning at least \$42k a year from home sharing. This is because those high-earning hosts receive half of the revenue and booked nights overall. Such statistical biases accurately depict the Airbnb experience, since again the revenue distribution for listings (as observed by an auditor for instance) are not the same as the distribution experienced by travelers themselves.

As a final note, and in light of New York's recent ban on advertising entire homes for stays shorter than 30 days [13], we present an early result of using our method to study the effect of new and proposed regulations. According to our estimates, banning stays under 30 days would result in entire home renters (in theory the only type affected) losing more than 85% of their revenue. We have not yet estimated

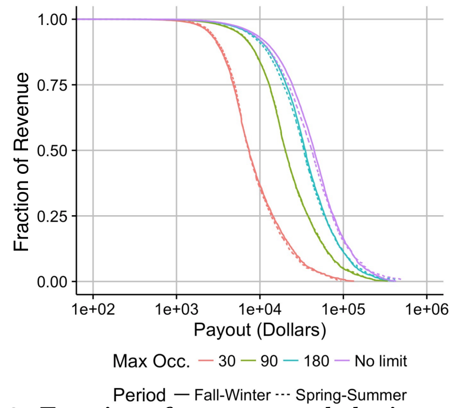

Figure 8: Fraction of revenue made by income levels: for every x, y is the fraction of total revenue going to hosts earning at least x.

the characteristics of affected listings, but our methodology will help in the future to study that effect. Multiple cities have considered a different policy: limiting the total number of nights on which a listing may be booked in a year. Using our method, we recomputed the revenue earned by each host during those periods when the occupation ratio is by law capped at a maximum[2]. The results are shown in the previously mentioned figures with different colors (Figures 7 and 8). We observe that enforcing 180 days a year as a maximum has little effect, except on a handful of hosts. Enforcing a maximum occupancy of 90 days a year appears as an interesting middle ground: The majority of hosts using the platform as extra income to pay the rent (for instance those earning less than \$10k) see almost no revenue reduction (the median host remains above \$8k). On the other hand, high earners are especially affected (with the top 10% dropping from \$50k to \$25k). Given that using dedicated homes for Airbnb often means paying full rent, this legislation can make illegal hotels economically unsustainable. Note however that such a law would greatly affect the supply of listings on Airbnb (remember that Figure 5 suggests that 80% of the revenue comes from listings used at least 90 days a year). This could also increase the nightly price of available places, and may not be popular among travelers. A more drastic regulation enforcing a maximum of 30 days a year would entirely transform the income of almost all Airbnb hosts and indeed all its associated economy (listings with more than 30 days a year account for 95% of the revenue for any period and listing types). While this might be a last resort as far as preventing home sharing abuse is concerned, it is just this kind of social trade-off that must be heavily debated before being put into practice. Beyond the scope of this paper, our data sets and methodology will be provided for the research community to conduct a richer analysis and study the effect of such regulation on different neighborhoods.

7. CONCLUSIONS

In this paper, we show the feasibility and advantage of frequent measurements to obtain scientifically reproducible

[2]We scaled down the revenue proportionally to this cap whenever occupancy was higher. It is a crude estimate that does not consider seasonal price variation and other effects.

and accurate data about both the supply and demand side of Airbnb's marketplace. This information already proves useful for informing the current debate surrounding the regulation of home sharing platforms. As our separate evaluation suggests, previously reported – and heavily debated – facts, while strictly speaking correct, are used to draw conclusions that oversimplify how the sharing economy works. For instance, claims made by Airbnb using a "median" listing ignore the inspection paradox, and a much different picture emerges once the full distribution is available. Indeed, there is strong evidence that a large fraction of booked stays, and thus revenue, goes to a small number of listings and hosts.

Multiple other controversial issues could benefit from this approach, such as the benefits and drawbacks of short term accommodation on the spread of tourism or the scarcity of affordable housing. We hope that our data and techniques, which will be shared with the research community, will facilitate this important debate. Beyond this paper, those methods and observations might generalize to the study of other economies with dynamic availability, such as the growing online market for cleaning or on-demand labor.

8. REFERENCES

[1] Airbnb. Data on the Airbnb community in New York City . Technical report, AirBnB corporation, Dec. 2015.

[2] Aribnb. Airbnb economic impact. `http://blog.airbnb.com/economic-impact-airbnb/`.

[3] A. C. Brian Chesky. Quora: What percent of airbnb hosts leave reviews for their guests? `https://www.quora.com/What-percent-of-Airbnb-hosts-leave-reviews-for-their-guests`, 2012.

[4] Budget and L. Analyst. City and county of san fransisco board of supervisors budget and legislative analyst - policy analysis report. `http://sfbos.org/sites/default/files/FileCenter/Documents/52601-BLA.ShortTermRentals.051315.pdf`, 2015.

[5] M. Cox. NYC: Battle Against Airbnb Hosts with Multiple Entire Home Listings Won, but the War Against Commercial Listings Continues . Technical report, insideairbnb association, July 2016.

[6] M. Cox and T. Slee. How Airbnb's data hid the facts in New York City . Technical report, insideairbnb association, Feb. 2016.

[7] B. G. Edelman, M. Luca, and D. Svirsky. Racial Discrimination in the Sharing Economy: Evidence from a Field Experiment. *SSRN Electronic Journal*, 2015.

[8] B. Efron. Censored data and the bootstrap. *Journal of the American Statistical Association*, 76(374):312–319, 1981.

[9] C. Elliot. Airbnb Runs 'Illegal Hotels,' Hotel Industry Study Claims. *Fortune*, pages 1–3, Jan. 2016.

[10] M. D. Ernst et al. Permutation methods: a basis for exact inference. *Statistical Science*, 19(4):676–685, 2004.

[11] A. FAW. Smart pricing. `https://www.airbnb.com/help/article/1168/how-do-i-turn-smart-pricing-on-or-off`.

[12] A. Fradkin, E. Grewal, D. Holtz, and M. Pearson. Bias and reciprocity in online reviews: Evidence from field experiments on airbnb. In *Proceedings of the Sixteenth ACM Conference on Economics and Computation*, EC '15, pages 641–641, New York, NY, USA, 2015. ACM.

[13] A. J. . Hawkins. Airbnb is now banned from listing short-term rentals in new york. `http://www.theverge.com/2016/10/21/13361536/airbnb-new-york-cuomo-bill-ban-short-term`.

[14] J. W. ONeil and Y. Ouyang. From Air Mattresses to Unregulated Business: An Analysis of the Other Side of Airbnb. Technical report, Penn State University, Jan. 2016.

[15] G. Quattrone, D. Proserpio, D. Quercia, L. Capra, and M. Musolesi. Who Benefits from the Sharing Economy of Airbnb. *WWW '16: Proceedings of the 25th international conference on World Wide Web*, Apr. 2016.

[16] E. T. Schneiderman. Airbnb in the city. Technical report, New York State Office of the Attorney General, 2014.

[17] D. Streitfeld. Airbnb Listings Mostly Illegal, New York State Contends. *New York Times*, Oct. 2014.

[18] A. Stulberg. Airbnb's multi-unit hosts and "commercial listings" account for growing share of NYC business. *New York Real Estate News*, pages 1–2, July 2016.

[19] G. Zervas, D. Proserpio, and J. Byers. The rise of the sharing economy: Estimating the impact of Airbnb on the hotel industry. Technical report, Boston U. School of Management, 2013.

[20] G. Zervas, D. Proserpio, and J. Byers. The Impact of the Sharing Economy on the Hotel Industry: Evidence from Airbnb's Entry in Texas. *EC '15: Proceedings of the sixteenth ACM conference on Economics and computation*, pages 1–36, Jan. 2014.

[21] G. Zervas, D. Proserpio, and J. Byers. A First Look at Online Reputation on Airbnb Where Every Stay is Above Average. Technical report, Boston University, 2015.