Airbnb's Business and Arguments about Data: *Address to the Asper Review of International Business and Trade Law*

April 23, 2018 – Winnipeg, Manitoba

T O M   S L E E [*]

SPEECH NOTES AND FIGURES

"Sound legal and policy decisions regarding the Sharing Economy require a clear understanding of Sharing Economy business models. Such an understanding would normally be based on statistical facts, but when it comes to Airbnb, even those have been the subject of continuing debates: Airbnb releases aggregated data only, and the company has disputed the accuracy of more fine-grained "scraped data" used by critics.

This talk paints a picture of Airbnb's business based on "scraped data," collected by the author from the company's website.[1] It shows that, despite the debates, there is in general good agreement between the data Airbnb has presented and the data collected from the web site by independent researchers. The apparent disagreements come almost entirely from differences in how the data is presented.

---

[*] Tom Slee is the author of the book "What's Yours is Mine: Against the Sharing Economy" (OR Books, 2015), an early study of the Sharing Economy which has been translated into five languages. After gaining a PhD in Theoretical Chemistry from McMaster University and carrying out postdoctoral work at the universities of Oxford and Waterloo, he has had a long career in the software industry. His Airbnb data sets and analyses have been used in numerous academic papers and have been the source of stories in the New York Times, The Guardian, The Economist and more.
[1] All the data used in this presentation is available for download from the author's web site, at <tomslee.net/airbnb-data-collection-get-the-data>. The data was collected from the Airbnb public web site using code that is publicly available from <github.com/tomslee/airbnb-data-collection/>.

<h3 align="center">TABLE OF CONTENTS</h3>

I.   INTRODUCTION

The sharing economy has arrived with its fair share of legal and policy controversy, centering on the two leading companies: Uber and Airbnb. A prerequisite for sensible policy and legal decisions is an agreement concerning what kind of business these companies are, but over the last several years even that seemingly-elementary question has been repeatedly disputed. In a phrasing repeated frequently by the company, Airbnb's Global Head of Public Policy and Public Affairs Chris Lehane says the platform enables "everyday people to share their homes across Canada in order to help make some extra income to help make ends meet."[2] Meanwhile, Thorben Weiditz of Fairbnb, a Toronto-based organization that has called for increased regulation of Airbnb's business, presents a different picture: "Most of Airbnb's short-term rental business takes place in what the company calls Toronto's 'conventional hotel market'."[3] So which is it?

How much do hosts earn? How much of Airbnb's business comes from "commercial" hosts? These would seem to be quantitative questions that could be answered by straightforward data from the company, anonymized if necessary. This is one reason why Teresa Scassa recently argued for a public interest role in reliable and comprehensive data concerning the Sharing

---

[2] CNW Group/Airbnb, "Airbnb names new Canadian executive", *Newswire,* (25 July 2016), online: <newswire.ca/news-releases/airbnb-names-new-canadian-executive-588138212.html>.

[3] Thorben Wieditz, Fairbnb.ca Coalition, "Squeezed Out: Airbnb's Commercialization of Home-Sharing in Toronto" (March 2017), online: <researchgate.net/publication/319207035/download>.

Economy.[4] But the data-driven part of the policy debate has been just as heated as disputes over other aspects of the sharing economy. Airbnb itself and its critics have repeatedly presented data-driven portraits of the company and its business. What is surprising is that the pictures are so different.

Airbnb itself does not release listing-level data for any city.[5] There is no company-provided spreadsheet or database with an entry for each of the approximately 15,000 rooms, apartments, and houses being offered for rent in Toronto, containing information such as the approximate location, the number of nights it has been occupied, and an identifier for the host. Instead Airbnb limits itself to aggregating the data before release, providing summary statements and tables in its reports. This pre-aggregation means that people outside Airbnb cannot use the data to ask their own questions but are limited to Airbnb's presentation. For example, in a 2016 report on "The Airbnb Community in Ontario",[6] Airbnb states that "88 percent of hosts have one entire home listing",[7] but say nothing about the number of listings those other 12 percent have.

Independently-collected data provides an alternative source. It is obtained by "scraping" the Airbnb public web site. That is, running a program that goes repeatedly to the Airbnb public web site, runs the same kind of searches that human users run, and analyzes the web pages that the searches return to pick out the key items. I have been running this kind of data collection since 2013, and the data I have collected has been used by journalists, academics, and public interest groups investigating the Airbnb phenomenon. In the recent debate over Airbnb regulation in Toronto my data was used by the Canadian Centre for Policy Alternatives[8] and by Fairbnb.ca[9] in putting together their reports and policy recommendations.

There are others who carry out data scraping, in many ways providing higher quality data than my own. Inside Airbnb is a one-person effort that

---

[4] See generally Teresa Scassa, "Sharing Data in the Platform Economy: A Public Interest Argument for Access to Platform Data" (2017) 50:4 UBC L Rev 1017 at 1017–1071.

[5] Colin Horgan, "Airbnb's Toronto Data Is a Confusing Mess", *Maclean's* (8 September 2016), online: <macleans.ca/news/canada/airbnbs-toronto-data-is-a-confusing-mess/>.

[6] Airbnb, "The Airbnb Community in Ontario" (7 September 2016), online (blog): *Airbnbcitizen* <airbnbcitizen.com/the-airbnb-community-in-ontario/> [Airbnb, "Ontario"].

[7] *Ibid.*

[8] Zohra Jamasi & Trish Hennessy, Canadian Centre for Policy Alternatives, "Nobody's Business" (22 September 2016) at 24, online: <policyalternatives.ca/airbnb>.

[9] See Thorben Weiditz, Fairbnb.ca Coalition, "Squeezed Out: Airbnb's Commercialization of Home-Sharing in Toronto" (29 February 2017), online: <fairbnb.ca/wpcontent/uploads/2017/09/Fairbnb_Report_Feb_29.pdf>.

provides "an independent, non-commercial set of tools and data that allows you to explore how Airbnb is really being used in cities around the world."[10] AirDNA is a company that has made a business of selling Airbnb data & insights both to prospective hosts (who may wish to use it to set prices, for example) and also to academics and cities.[11] In addition, some academics have written their own code and done their own data collection.

Airbnb is dismissive of scraped data. Here is an excerpt from an Airbnb email to the City of Toronto:

> 'Scraped data' is inherently unreliable, as it relies upon extracting a limited set of data from our public-facing platform, then using this data to build models that make predictions and conclusions about the entirety of the home sharing community. These erroneous findings are a major contributor to the myths about home sharing.[12]

Here I review what the data I have collected from Airbnb's site tells us, and show that, surprisingly given the company's statements, the broad outlines are consistent with Airbnb's own reports. The differences are not so much in the data itself as in its presentation. Readers are invited to make up their own mind about which presentation is more relevant to public policy debates.

II. TOTAL LISTINGS

We start with an uncontroversial statistic: what is the total number of listings in a city?[13]

---

[10] See generally Inside Airbnb, "About Inside Airbnb", online: <insideairbnb.com/about.html>.

[11] See generally AirDNA, online: <airdna.co/>.

[12] This is contained in an email from Airbnb to the City of Toronto. This email is not available to the public, but the same text is present in an email from Airbnb to the City of Edmonton. See Alex Dagg, Public Policy Manager, Airbnb Canada, Letter to Colton Kirsop, Senier Planner, Development and Zoning Services (25 January 2018), 1 at n 1, online: <sirepub.edmonton.ca/sirepub/cache/2/xus21t2fwthoazxi4e0z1u3d/71515709032018072204 707.PDF >.

[13] The term "city" is used for convenience to describe an area investigated by a scrape or "survey", even though some surveys cover larger regions or whole countries. While each separately rentable space is sometimes still called a "room" on Airbnb, we shall see that this terminology is no longer appropriate, and so the general term "listing" is used to indicate a space provided for rent.

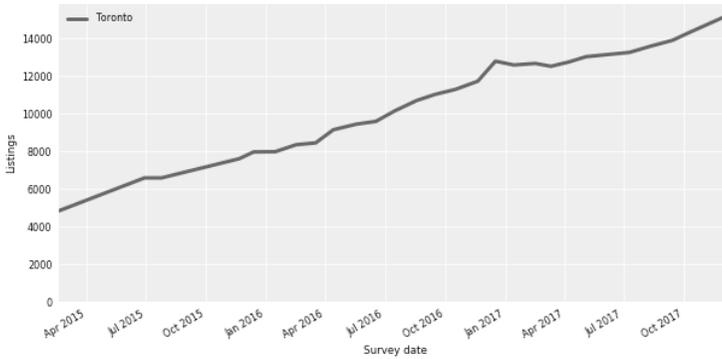Here in *Figure 1*, Toronto is provided as an example:

Figure 1 shows the number of listings in Toronto over the years 2015–2017, showing that it has increased steadily from around 5,000 in early 2015 to over 15,000 at the end of 2017. The figure matches that reported in Maclean's magazine[14] and that produced by Inside Airbnb. The total number of distinct hosts in late 2016 is between 8,000 and 8,500, which matches well with Airbnb's own assessment of 8,600 hosts in Toronto at some unspecified time in 2016.[15] The agreement verifies that independent surveys are locating most of the listings in a city, and correctly identifying the host ID values. Next, in *Figure 2*, the total listings for a selection of other Canadian cities and towns are represented.[16]

---
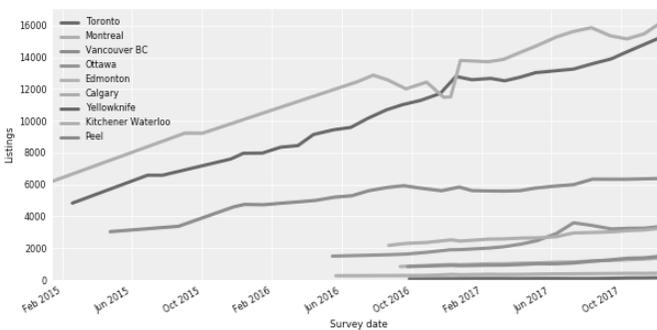
[14] Colin Horgan, "Airbnb's Toronto Data Is a Confusing Mess" *Maclean's*, September 2016, online: <macleans.ca/news/canada/airbnbs-toronto-data-is-a-confusing-mess/>.

[15] Airbnb, "Ontario", *supra* note 6.

[16] Steady growth can be seen, as well as variations in the data due to potential listing fluctuations or survey mistakes.

Indeed, the following figure – *Figure 3*, shows that the data collection is good enough to capture some exceptional events. The spike in listings in Washington DC coincides with Donald Trump's inauguration (we cannot tell if the listings were rented out, and if so whether to supporters or opponents), and the spike in Ottawa coincides with Canada's 150th birthday celebrations. Across the world, the pattern is one of steadily increasing listing counts.
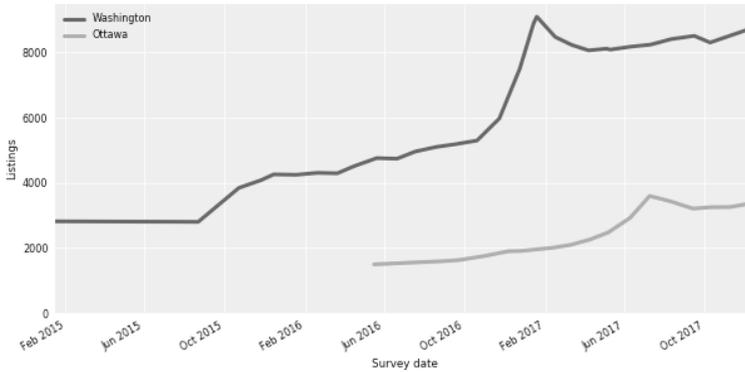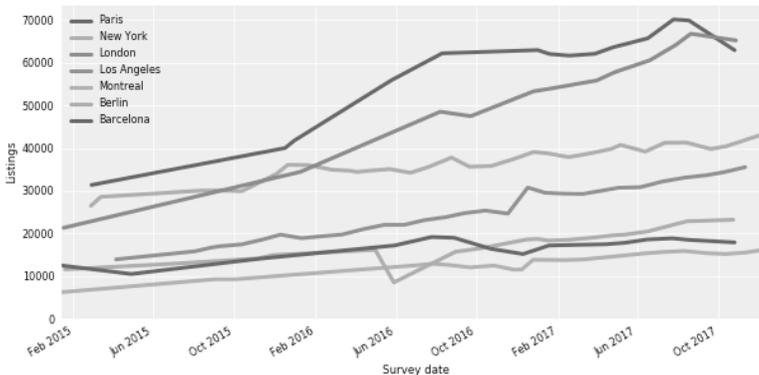

Figure 4, below, shows data taken from several major tourist cities. The one major drop – in Berlin in mid-2016 – is the result of a policy change where residents were not permitted to list entire homes on the site. The policy was not strongly enforced, and after a short break many listings reappeared.

The overall listing count verifies that independent scrapes capture useful and accurate information about the number of Airbnb listings, as well as their basic characteristics, such as hosts. It helps set the ground for policy discussions when we know, for example, that the number of Airbnb listings in Toronto

(15,000) is similar to the number of hotel rooms used by tourists (Toronto has an estimated 40,000 hotel rooms, with two-thirds of hotel traffic being used for business travel. Despite repeated efforts, Airbnb has not made big inroads in the business travel market).

III. LISTING TYPES

A second necessary set of data points is the relative proportions of different listing types. Listings on Airbnb are classified as one of three "room types," a phrase that is used in the technical codes on the website, but now outdated.[17]

One type is the "shared room", which corresponds to the original Airbnb story, when the founders of the company were newly graduated design students struggling to pay the rent. Seeing a conference coming to town, they rented out airbeds in their apartment as an "Airbed and Breakfast". The rest is history, and a mockup of the original apartment has been recreated in the company's headquarters.

A second listing type is the "private room," which is what many people think of when they think of Airbnb: it corresponds to renting out a spare room.[18]

The third type is the "entire home/apartment," which means that a guest has sole use of a complete living space.[19] It is immediately obvious that the original "shared room" plays almost no part in the company's business model today, comprising well under 5% of listings on the site. Despite the way Airbnb is commonly described, "private room" listings make up only about a third of the total, and "entire home/apartment" listings are the majority on the site. The proportion of each stayed relatively constant over two years of study. The following representation, in *Figure 5*, shows the relative proportion of different listing types in Toronto:

---

[17] "What do the different home types mean" (last visited 7 September 2018), online: *Airbnb* <airbnb.ca/help/article/317/what-do-the-different-home-types-mean>.
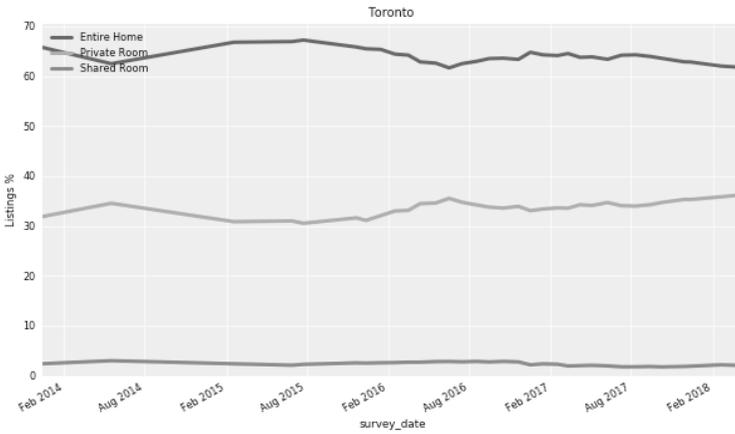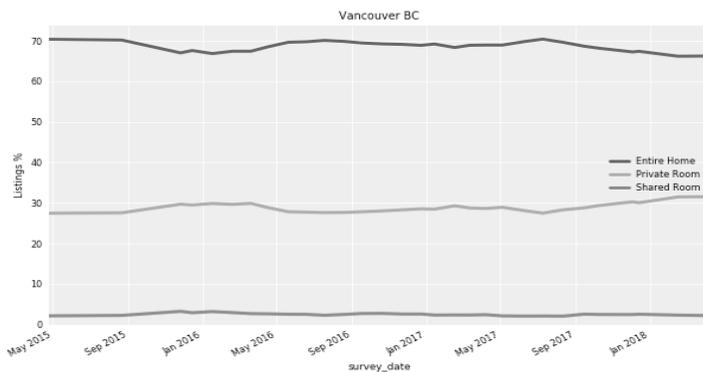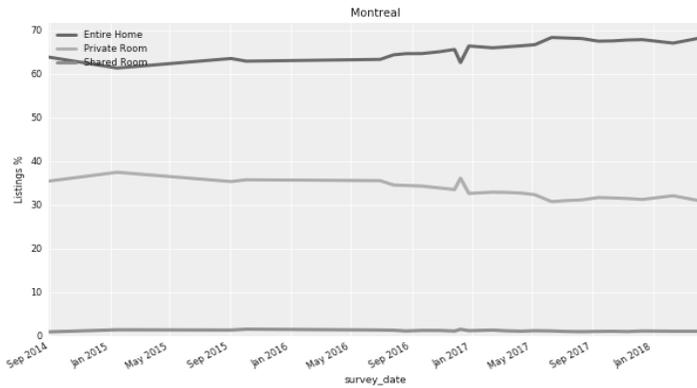[18] *Ibid.*
[19] *Ibid.*

Toronto

Figure 6 and Figure 7, following, show the same data sets for Montreal and Vancouver, respectively. Each show that the bulk of the business is "Entire home" rentals.

Montreal

Vancouver BC

The following bar chart, *Figure 8*, shows the relative proportion of different listing types for a selection of Canadian cities, in late 2017. Smaller cities such as Yellowknife have a much bigger representation of private rooms, but the big tourist destinations are dominated, as we have seen, by entire home listings.
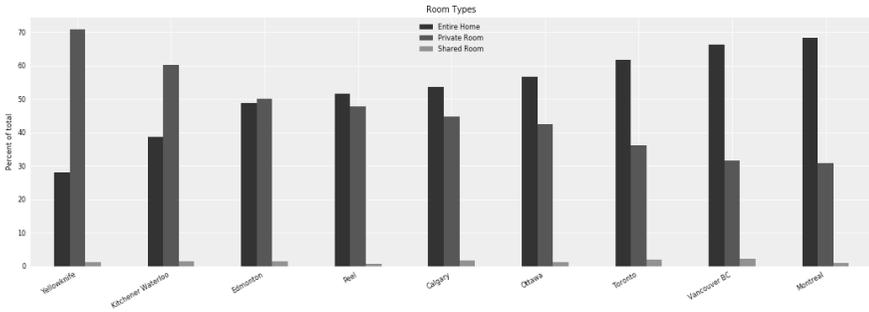

Figure 9, following, shows a comparison of cities around the world. The picture is varied, and the reasons for the variation in listing types would require a knowledge of the conditions and history of each location. It is left as an exercise for the reader to investigate such.

For Canada, the picture is clear that in the big cities – which make up the bulk of tourist travel destinations in the country, a sizeable majority of Airbnb listings is of the "entire home/apartment" type. Regulation of Airbnb should not, therefore, proceed as if it were built on a business model of renting out the "spare room".
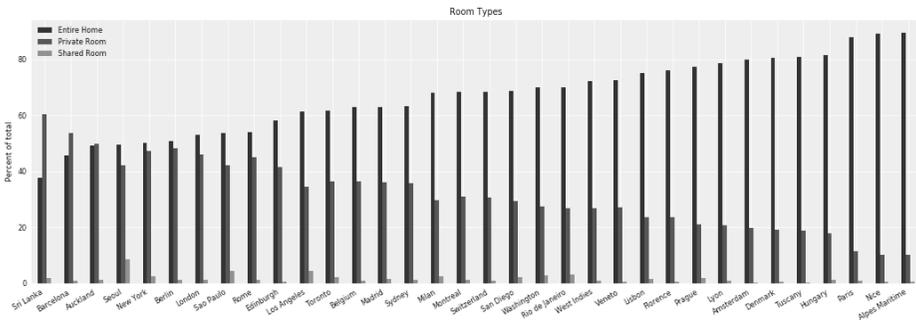
IV. MULTIPLE OWNERSHIP

It is when we turn to multiple ownership that the controversy starts. Here is one lens, under which it seems like there has been a big difference between Airbnb's view of the world and that of its critics.

Airbnb statements about the incidence of multiple ownership in their major cities are consistent, emphasizing the high percentage of home-sharing hosts, as opposed to the more commercial hosts with multiple listings:

- "87 percent of Airbnb hosts rent out the home they live in" (New York, 2014)[20]
- "95% of Airbnb hosts who share an entire home have only one entire home listing" (New York, 2018)[21]
- "87 percent of hosts rent the homes they live in" (Amsterdam, 2014)[22]
- "88% of Airbnb hosts in Washington, D.C. are sharing space in their permanent home" (Washington 2016)[23]
- "As many as 88 per cent of Airbnb hosts in the capital [Dublin] said they shared their primary residence" in 2016[24]
- "73 percent of hosts share just one listing" Barcelona 2016[25]

Meanwhile, critics convey a different message: for example, a 2016 report on the Airbnb market in New York City, conducted by Murray Cox and myself, investigated an Inside Airbnb claim that "[t]he percentage of traffic going to multiple-listers has been a constant 30% of the total during all 2015."[26] At first glance, it seems that this and the above-mentioned figures are different, but a moment's thought shows that they could in fact be consistent.

---

[20] Airbnb, "New Study: Airbnb Generated $632 Million in Economic Activity in New York", *Airbnb,* online: <airbnb.ca/press/news/new-study-airbnb-generated-632-million-in-economic-activity-in-new-york>.

[21] Airbnb, "One Host, One Home: New York City" (February 2018), online (pdf): <https://2sqy5r1jf93u30kwzc1smfqt-wpengine.netdna-ssl.com/wp-content/uploads/2018/03/One-Host-One-Home-_-February-2018-Update.pdf.>.

[22]"Airbnb Economic Impact" (last visited 7 September 2018), online (blog): *Airbnb* <blog.atairbnb.com/economic-impact-airbnb/>.

[23] Airbnb, "Airbnb, Housing, and the District of Columbia" (last visited 7 September 2018), online (pdf): *Airbnbcitizen* <airbnbcitizen.com/wpcontent/uploads/sites/24/2017/04/2017-04-Airbnb-D.C.-Housing-Report.pdf>.

[24] Charlie Taylor, "How much did Airbnb hosts in Dublin earn last year?" *Irish Times* (26 Jan 2017), online: <irishtimes.com/business/economy/how-much-did-airbnb-hosts-in-dublin-earn-last-year-1.2951124>.

[25] Airbnb, "Airbnb community boosts Barcelona economy by €740 million", *AirbnbCitizen* (16 March 2016), online: < airbnbcitizen.com/airbnb-community-boosts-barcelona-economy-by-e740-million/>.

[26] Murray Cox & Tom Slee, "How Airbnb's data hid the facts in New York City" (10 February 2016), online: <whimsley.s3.amazonaws.com/wordpress/wp-content/uploads/2016/02/how-airbnbs-data-hid-the-facts-in-new-york-city.pdf>.

Consider a small Airbnb "city" with just four hosts and six listings:
- Aslam, Rose, and Alex each have a single listing, renting out their own homes.
- Sophie is running a small Airbnb hosting business and has 3 listings on the site.

For this city, both the following statements are true:
1. Seventy five percent of hosts "rent out the home in which they live".
2. Commercial offerings make up half of the listings in the city.

Statement 1 is Airbnb's favoured presentation of the distribution, while critics emphasize the second. For most public and community purposes it is the second data point that matters – the percentage of listings, not the percentage of hosts. A visitor to the city browsing the Airbnb web site will see six listings, half of which are commercial offerings, and a city government concerned about what is happening on the ground cares about the percentage of commercial listings, not the percentage of hosts.

The data set of *Figure 10*, below, shows that 88% of hosts rent out a single listing, which is the same percentage as reported by Airbnb.[27]
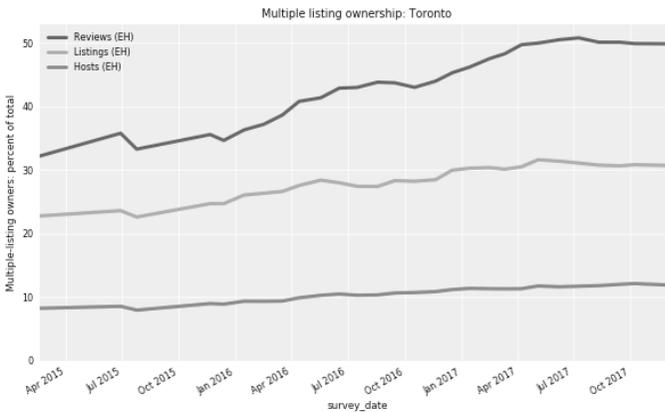

This close match again shows that independently-collected data present a realistic picture of the Airbnb business.

The percentage of listings belonging to the 12% of hosts with more than one listing is more like 30%. Airbnb is silent on this question, but there is no contradiction in the data between Airbnb's aggregate statements and those we

---

[27] Airbnb, "Ontario Community", *supra* note 6 at 3.

see from independent data after all. The differences are simply in the presentation. The top line in *Figure 10*, below, is an estimate of revenue for single and multiple listers, based on overall review counts (see also below) and shows an even higher figure, with perhaps half of Airbnb's revenue in Toronto coming from "commercial" listings.

    Figure 11 and *Figure 12*, below, show the same data for Montreal and Vancouver, respectively. The pattern is similar, as is the agreement with Airbnb's claims. The percentage of commercial hosts is relatively small, at around 10% of the total, and matches Airbnb's statements as well. The percentage of listings belonging to commercial hosts is higher, and Airbnb has been silent on this important policy data point.
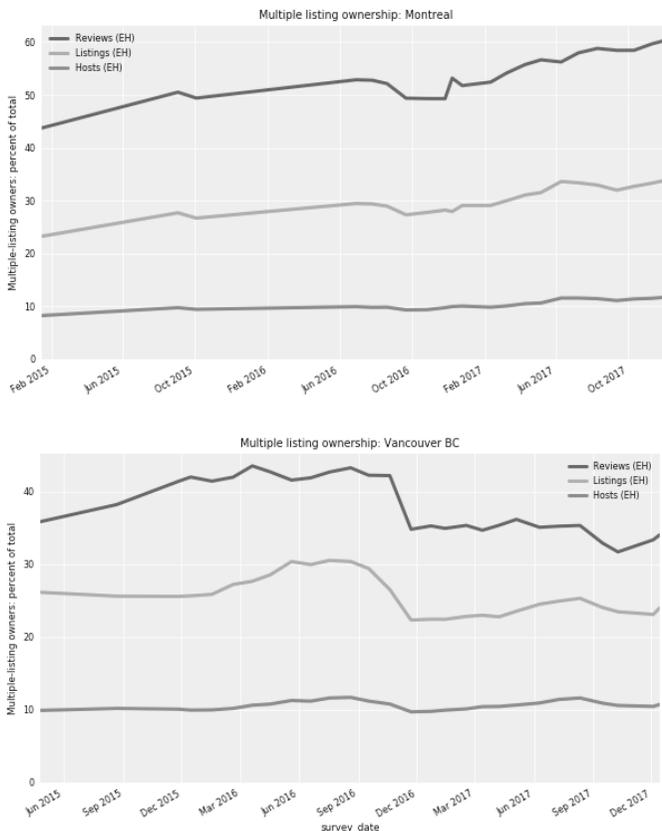

    Using the total number of reviews for each listing as a proxy for the total visits to the listings, the commercial listings are busier than those owned by casual hosts. Perhaps this is not surprising: a commercial owner has an

incentive to seek out popular locations and to rent out the listing as often as possible, while a casual home-sharer can rent out only their own home, no matter whether it is in a popular location or not. As a result, it seems that multiple-listing hosts make up about half of Airbnb's total business in Canada's three largest cities.

Figure 13, below, shows an estimate of the percentage of multiple listings in a number of cities and regions around the world. The diversity is striking, with some cities being much closer to a single-listing population, while others are dominated by hosts with multiple listings. Again, interpretation of these values is left as an exercise for the reader.

For Canadian purposes, the conclusion is that independently collected data is consistent with that provided by Airbnb, but that it also answers more policy-relevant questions, by providing a breakdown of "own home" vs multi-lister "commercial" listings by listing, and not just by host. Airbnb's own data remains silent on this important policy question.
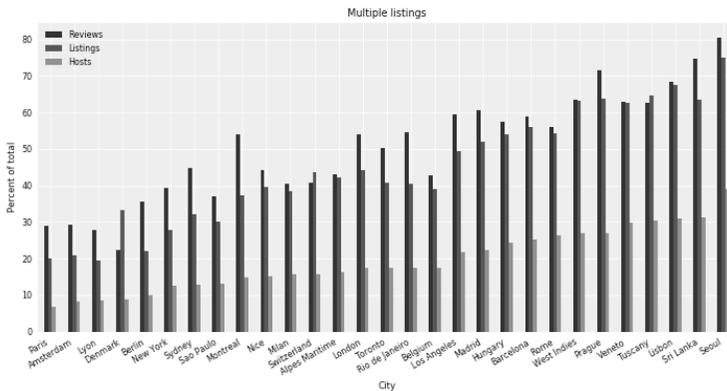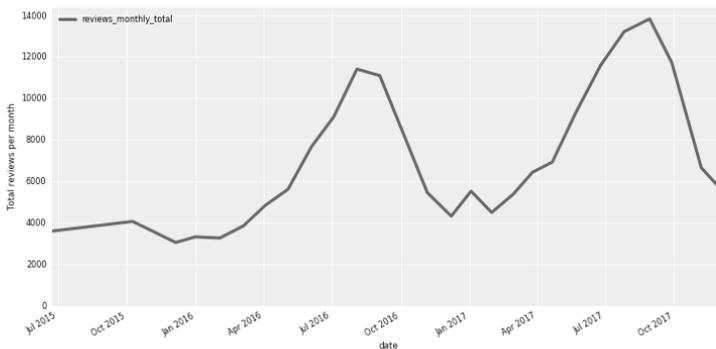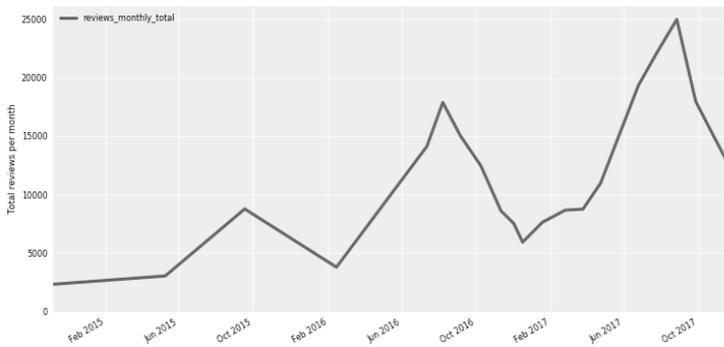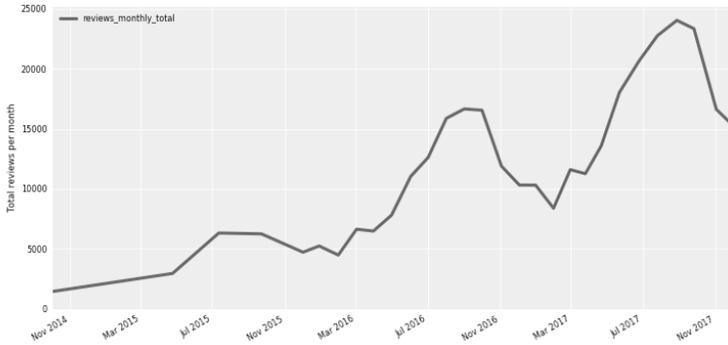

V. Occupancy

While independent data on multiple listing ownership has been reliable, data on occupancy is more uncertain. There is no publicly-accessible data that state whether a listing is rented out on a particular night. Rented listings will be marked as "unavailable" on a host's public calendar, but a listing may also be marked as unavailable for other reasons: the host may not be making it available to the public on that day.

While AirDNA bases its estimates on calendar entries, I have used an even more approximate method of using review counts as a proxy for relative numbers of visitors.

Provided with a large number of listings, an idea of the broad trends should be apparent. For example, the following figures – *Figure 14*, *Figure 15*, and *Figure 16* – show that using reviews as a proxy for visits (in Toronto, Montreal, and Vancouver, respectively) captures the expected seasonal fluctuations in travel.

Again, the picture of occupancy painted by Airbnb differs from that of the critics. Bear in mind that Airbnb emphasizes that listings are only occupied occasionally, with typical hosts renting for maybe one night in ten:

- **Ontario:** "typical hosts are earning $3,900 annually from renting out their own primary residence for three to four nights a month" (Alex Dagg, 2016)[28]
- **New York:** "the typical host earns $7,530 per year" (Airbnb Economic Impact)[29]
- **New Zealand:** "the typical Airbnb host earns NZD$80 per week or NZD$4,200 a year" (Airbnb, 2018)[30]

Most of these listings are only shared occasionally – nearly 60% of "entire home" listings are rented less than 90 days per year. (Washington DC)[31]

A small-city example is useful, again, to make sense of how these numbers are presented. Consider the case of four hosts, who each have a single listing, but while three (Aslam, Rose, Alex) only rent occasionally, Sophie rents intensively.

- Aslam's listing is occupied for 0 nights per year
- Rose's listing is occupied for 30 nights
- Alex's listing is occupied for 42 nights
- Sophie's listing is occupied for 180 nights

The following statements are both true:
1. The typical host rents for only 36 nights in the year. The "typical host" is the median value, which for a population of four is midway between Rose and Alex.
2. Most of the guest traffic (i.e. over two-thirds) is via Sophie's "ghost hotel."

---

[28] Sammy Hudes, "Typical Airbnb hosts in Canada earn $3,900 a year", *The Hamilton Spectator*, (7 September 2016), online: <thespec.com/news-story/6845187-typical-airbnb-hosts-in-canada-earn-3-900-a-year/>.

[29] Airbnb, "Airbnb Economic Impact" (accessed 30 August 2018), *atairbnb* (blog), online: <blog.atairbnb.com/economic-impact-airbnb/>.

[30] Airbnb, "Airbnb Welcomes More than 1.4m Guest in 2017" (2 March 2018), online: *SCOOP Business* <scoop.co.nz/stories/BU1803/S00053/airbnb-welcomes-more-than-14m-guests-in-2017.htm>.

[31] Airbnb, "Airbnb, Housing, and the District of Columbia"(20 April 2017), online (pdf): *Airbnbcitizen* <airbnbcitizen.com/wp-content/uploads/sites/24/2017/04/2017-04-Airbnb-D.C.-Housing-Report.pdf>.

Again, the second statement is the important one for most public policy purposes and is the figure that will matter to neighbourhood residents. Again, Airbnb is focusing on the less relevant data points in its own presentation. Understandably, condo owners will be concerned with the number of commercial visitors in their building and not with the number of listings on a website. Aslam's listing, which is never rented, is of no concern to anybody (is a host with no guests still a host?) and discussing "typical hosts" gives an unjustifiably high importance to low occupancy hosts.

In many cities, a limit of 90 days has been proposed as a reasonable upper limit for the kind of casual business that Airbnb claims to be involved in. This time, we actually have some data from Airbnb, courtesy of a freedom of information request by Fairbnb in Toronto.

Figure 17, to the side, shows that about 60% of listings rent out for less than 90 days a year, and the median value is about 60 nights, which matches Airbnb's statements about "typical" hosts.

But reworking that again: where are most nights spent? A few hosts with heavily-occupied listings can count for a large proportion of the nights spent in Airbnb listings. Below, *Figure 18* shows the data points from *Figure 17* from two perspectives. On the bottom-left is an interpolation of the values given, which confirms the typical host occupancy (the red line) at 60 days. The figure on the right is the cumulative occupancy. It shows that well over 80% of the nights spent in Airbnb listings are spent in places rented out over 60 days a year, and about 75% of nights are spent in listings rented out over 90 days a year.

Guest information
Average age of booking guests: 33
Guest origin:
• Europe: 14%   • North America: 77%   • Asia: 4%   • Australia: 2%   • Latin America: 3%   • Africa: 1%
Nights hosted for a typical listing: 55
Nights hosted per listing -- all listings:
• Under 30 nights: 35%   • 21-60 nights: 16%   • 61-90 nights: 10%   • 91-120 nights: 8%   • 121-180 nights: 11%   • More than 180 nights: 19%
Currently active listings by type:
• Entire home: 64%   • Private space: 33%   • Shared room: 2%

Again, the data from critics and Airbnb is broadly comparable, despite the surface differences, but the presentation by Airbnb is a misdirection.

VI. CONCLUSIONS

My goal in this talk has been to clarify some of the disputes around Airbnb in Canada's cities. One thread has been that, despite Airbnb's protestations,

scraped data can give a reasonable picture of Airbnb's business, albeit with gaps on the topic of occupancy. The disagreements are more a matter of presentation than the underlying data quality: data can be viewed from multiple perspectives, and letting Airbnb decide which perspective to choose is not in the public interest.

A second goal is to show that, while there are many hosts on the Airbnb web site who fit the archetype of a casual host, occasionally renting out the home in which they live, the company's business is driven significantly by "commercial" hosts with multiple listings and by high-occupancy listings that are rented frequently. The picture of "ghost hotels" has a good deal of truth to it, and lawmakers and policy discussions should take this into account in their debates.