A Statistical View of The Right to Vancouverism: Social Reproduction Placemaking in the Revanchist City

Gian Carlo Diluvi*

Client: Cheryl-lee Madden

June 2021

1 Introduction

reference blog post and talk about the 2018 chs.

statistical questions:

Q1: how sound are statistical analyses in blog post?

Q2: can we expand that analysis to account for low-income women?

Q3: can we deep dive into vancouver?

2 Statistical Considerations of the 2018 CHS PUMF

In order to protect the privacy of survey respondents (i.e. preventing any respondent or household to be identified), data obtained in the CHS is modified in various ways. [Statistics Canada, 2021, Section 6] goes into detail about the safeguards used by Statistics Canada. These include decreasing the level of geographic detail; grouping answers into categories in questions that contain many answers; adding random noise to some quantitative variables; and rounding very small or large quantitative values, which normally correspond to extreme (i.e. rare) households.

The modified data, called the *public use microdata file* (PUMF), is then made public. Naturally, analyses based on the PUMF will differ from those carried out using the full master file of Statistics Canada due to the data modification process used to decrease disclosure risk. [Statistics Canada, 2021, Section 7] explains in detail all the limitations of analyses based on the PUMF. Succinctly, however, the PUMF should not be used to carry out statistical analyses. Rather, it should be used to conduct exploratory data analyses that might indicate which models are appropriate and possibly to obtain preliminary estimates of variables of interest.

Related to the first statistical question, the PUMF cannot be reliably used to measure variability, and it also does not include bootstrap weights. Hence, practically any statistical test would produce

*Department of Statistics, University of British Columbia. Contact: gian.diluvi@stat.ubc.ca.

invalid results. This is because, even if point estimates are not necessarily way off when compared with the values obtained from Statistics Canada's master file, there is no way to reliably measure the quality of each estimate. More so, the accuracy of point estimates depends on the number of variables that are being cross-tabulated. Again, there is no objective and reliable way to measure this accuracy.

Once preliminary results have been obtained and an appropriate model (or family of models) selected, the PUMF user guide Statistics Canada [2021] recommends requesting access to the CHS master files. Statistical analyses based on the full data files will be valid and, furthermore, can be fit with a greater level of detail (e.g. geographically).

As for the second statistical question, the PUMF contains the gender of the reference person, i.e. the one answering the survey. This means that accounting for gender using the PUMF would be difficult as there is no indication of the relationship between the reference person and the main provider of each household. It is possible that the master file does contain information about each member of the household, including gender, although the PUMF guide is not very clear about this.

In terms of the third statistical question, the PUMF only contains geographic information at the census metropolitan area (CMA), which for British Columbia (B.C.) corresponds to three categories: Vancouver, other large cities, and the rest of B.C. The confidential master file of Statistics Canada, however, does contain more detailed information that would allow a more granular analysis. Specifically, analyses that include neighborhoods within Vancouver can only be done based on the full master file.

Another point of interest is that of low-income families. The PUMF contains multiple variables related to the economical situation of each household, such as the annual income of each household. This information can be directly used to analyze (with the caveats mentioned in the previous paragraphs) the situation of low-income households. However, it should be mentioned that using poverty or low-income indices that depend on multiple variables might be unfeasible for multiple reasons. First, it is possible that not all the variables are available in the PUMF. Second, if they are, cross-tabulating multiple variables can decrease the accuracy of point estimates based on the PUMF, as was mentioned before.

Finally, it is not possile to use the 2018 CHS to analyze the impact that the COVID-19 pandemic has had on Canadian households due to the pandemic occurring three years after the census was carried out. However, it would be possible to use the 2021 census to study the impact of the pandemic once it is made public, with the same considerations as mentioned above.

3 Exploratory Data Analysis

In this section, we will address the second and third statistical questions, i.e. the landscape of forced moves for low-income women in Vancouver.

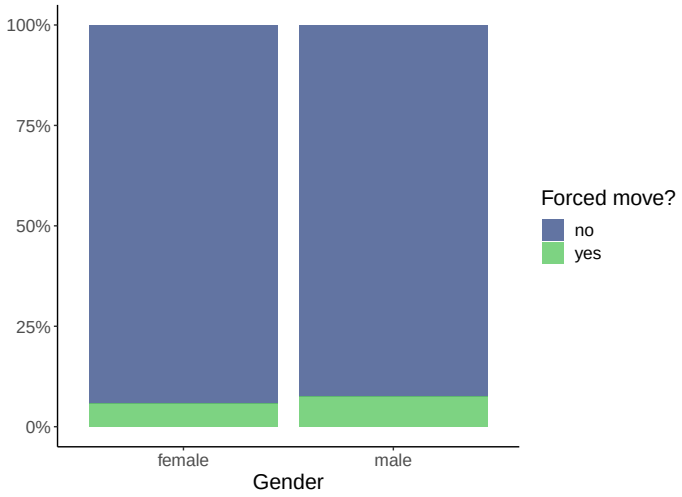

Figure 1: Percentage of people in B.C. that were forcefully moved from their house by gender.

As we discussed in the previous section, the PUMF contains the gender of the reference person. We will proceed assuming that the reference person corresponds to the primary provider of each household. Note, however, that this assumption might very well not hold, but it allows us to propose visualizations to understand the relationship between gender and forced moves. We suggest repeating these figures with the full master file of the 2018 CHS.

Figure 1 shows that slightly more men were forced to move out of their house, although it is not possible to assess whether the difference is significant in light of the previous section. Figure 2 shows that women that were *not* forced to move tend to have a higher income. Indeed, all the outlier points with really high incomes correspond to women who were not forced to move. On the other hand, the median income of women who were forced to move is smaller, although again it is not possible to test whether this difference is statistically significant (e.g. via a t-test or a permutation test).

4 Conclusion

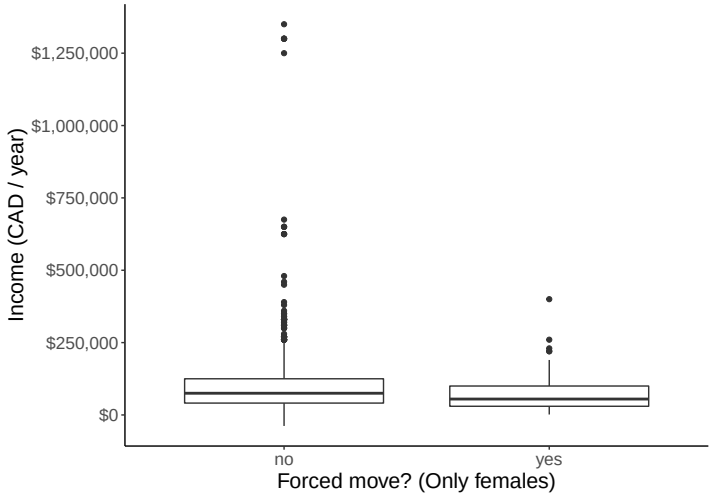

Figure 2: Distribution of income of women in B.C., divided by whether they were forced to move from their house.

References

Statistics Canada. 2018 Canadian Housing Survey: Public Use of Microdata File User Guide. In *Canadian Housing Survey, 2018*. Abacus Data Network, 2021.