

Better Household Living Conditions Improve Life Satisfaction

Rachel Oh, Yang Shang

19/10/2020

Abstract

This report describes the effects of household living conditions on life satisfaction. It is found that better household living conditions do improve life satisfaction to a certain degree. These results are important as it is fundamental that people know what makes them satisfied with their lives because no human would wish to have an unsatisfying life.

Introduction

In our studies, we established an interest about the relationship between people's life satisfaction and their living conditions. Resourced from Canadian GSS in 2017, our team is going to describe an individual's living conditions from the following aspects:

1. Household type: what type of residence is the respondent living in (house, apartment, etc.)?
2. Household size: how many residents (relatives and non-relatives), including the respondent, are living in the residence?
3. Ownership of the residence: is the respondent owning or renting the residence?
4. Cohabitation: Is the respondent living with a partner (married or not married)
5. Children in household: Is there any child (relatives and non-relatives) living in the residence? The detail of the above factors will be further described and explained in Section "Model".

Our prediction to the model is that: 1. people live in houses have higher life satisfaction; 2. the larger the household size, the higher the life satisfaction; 3. people owning the residence have higher life satisfaction; 4. people live with their partners have higher life satisfaction; 5. people live with children have higher life satisfaction.

Full code and data supporting this analysis is available at: <https://github.com/hyunbani/household-living-conditions-and-life-satisfaction>

Data

The data was collected from Statistics Canada's General Social Survey (GSS), which is an annual survey with statistical information on the living conditions and well-being of people living in Canada. In this report, we are using the data collected in 2017, through telephone interview.

People of interest were those who are 15 years of age and older, living in the 10 provinces. However, it would be costly and impractical to survey each and every household in Canada, so random sampling was chosen as a method to determine the sample of 43,000 people, which reflects the population. In order to ensure that the sample is an accurate reflection of the population as a whole, the survey results from all sampled households were collected, and only one eligible person per household was interviewed.

Since the answers are kept strictly confidential, the respondents must have provided the answers in a more honest manner, leading to an increase in accuracy of data information. However, the collection method was not very efficient since it took a lot of time to collect answers through telephone interview, and many people tend to not answer the phone calls.

People who did not answer the telephone calls were excluded from the sample. However, there still existed some non responses for some of the questions being asked, and these were recorded as ‘NA’ in the data. However, for better analysis, rows with answers including ‘NA’ are removed after reducing the data so that it only contains the variables of interest in this study.

The response variable ‘Life Satisfaction’ is measured on a scale of 0 to 10, with 10 being the highest satisfaction level and 0 being the lowest. The explanatory variables that explain household living conditions are included such as ‘Household Size’, ‘Owned’, ‘Partner’, ‘No Child’ and ‘Single House’.

The explanation of each explanatory variable is as follows:

- 1) ‘Household Size’ - the number of people each respondent lives with
- 2) ‘Owned’ - 1 if a respondent owns the house and 0 if he/she doesn’t
- 3) ‘Partner’ - 1 if a respondent lives with a partner and 0 if he/she doesn’t
- 4) ‘No Child’ - 1 if a respondent doesn’t have any child and 0 if he/she does
- 5) ‘Single House’ - 1 if respondent’s house type is single detached house and 0 if it’s not

Table 1 below shows the first ten rows of the data with the variables of interest.

Table 1: First 10 Rows of Raw Data

Case	Life Satisfaction	Household Size	Owned	Partner	No Child	Single House
1	8	1	1	0	1	0
2	10	2	1	0	1	1
3	8	2	1	0	1	1
4	10	2	1	0	1	0
5	8	2	0	1	1	0
6	9	2	1	0	1	1
7	4	1	0	0	1	0
8	10	1	0	0	1	0
9	8	1	1	0	1	0
10	5	6	1	0	1	1

Model

We are interested in explaining whether a person’s life satisfaction improves when he/she has better household living conditions, based on household size, house ownership, whether he/she has a partner, no child, and lives in a single detached house. In order to do so, we are going to fit a multiple linear regression model.

Multiple linear regression is a statistical technique that uses several explanatory variables to predict the outcome of a response variable. The goal of it is to model the linear relationship between the explanatory (independent) variables and response (dependent) variables.

As discussed in the data section, the dependent variable is Life Satisfaction and the explanatory variables are Household Size, Owned, Partner, No Child and Single House.

Formula and calculation of this model is as follows:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \beta_5 x_{i5} + \beta_6 x_{i6} + \epsilon$$

where, for $i = 20331$ observations:

y_i = dependent variable

x_i = explanatory variables
 β_0 = y-intercept (constant term)
 $\beta_1 \dots \beta_6$ = slope coefficients for each explanatory variable
 ϵ = the model's error term (also known as the residuals)

Results

Discussion

References

- R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- JJ Allaire and Yihui Xie and Jonathan McPherson and Javier Luraschi and Kevin Ushey and Aron Atkins and Hadley Wickham and Joe Cheng and Winston Chang and Richard Iannone (2020). rmarkdown: Dynamic Documents for R. R package version 2.3. URL <https://rmarkdown.rstudio.com>.
- Yihui Xie and J.J. Allaire and Garrett Golemund (2018). R Markdown: The Definitive Guide. Chapman and Hall/CRC. ISBN 9781138359338. URL <https://bookdown.org/yihui/rmarkdown>.
- Wickham et al., (2019). Welcome to the tidyverse. Journal of Open Source Software, 4(43), 1686. URL <https://doi.org/10.21105/joss.01686>
- Paul-Christian Bürkner (2017). brms: An R Package for Bayesian Multilevel Models Using Stan. Journal of Statistical Software, 80(1), 1-28. doi:10.18637/jss.v080.i01
- Paul-Christian Bürkner (2018). Advanced Bayesian Multilevel Modeling with the R Package brms. The R Journal, 10(1), 395-411. doi:10.32614/RJ-2018-017
- Canadian General Social Surveys (2017). General social survey on Family (cycle 31). URL <https://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&Id=335816>
- Stan Development Team (2020). RStan: the R interface to Stan. R package version 2.21.1. URL <http://mc-stan.org/>.
- Hadley Wickham and Evan Miller (2020). haven: Import and Export 'SPSS', 'Stata' and 'SAS' Files. R package version 2.3.1. URL <https://CRAN.R-project.org/package=haven>
- Jacob Kaplan (2020). fastDummies: Fast Creation of Dummy (Binary) Columns and Rows from Categorical Variables. R package version 1.6.2. URL <https://CRAN.R-project.org/package=fastDummies>
- Yihui Xie (2020). knitr: A General-Purpose Package for Dynamic Report Generation in R. R package version 1.30.
- Yihui Xie (2015) Dynamic Documents with R and knitr. 2nd edition. Chapman and Hall/CRC. ISBN 978-1498716963
- Yihui Xie (2014) knitr: A Comprehensive Tool for Reproducible Research in R. In Victoria Stodden, Friedrich Leisch and Roger D. Peng, editors, Implementing Reproducible Computational Research. Chapman and Hall/CRC. ISBN 978-1466561595