Does Perceived Sense of Belonging To Local Community Contribute to Life Satisfaction for Canadian-born Non-immigrants?

By Silas Kwok

This paper investigates self-reports of life satisfaction of Canadian-born non-immigrants in the 2018 Canadian Community Health Survey (CCHS), finding that a higher sense of belonging to local community is consistently associated with higher reports of general life satisfaction for this population.

I. Introduction

Life satisfaction is an important topic of research because it reflects people's quality of life and can be a powerful tool in terms of monitoring social progress and performing policy evaluation that goes beyond gross economic measures like output or profits (Veenhoven, 1996). Several factors that have been identified as contributors to life satisfaction include a sense of belonging to their communities, the amount of sleep they get, and the income they generate. Research finds that close relationships and personal contact are the greatest predictors of lifetime happiness (Solan, 2017) and that people who sleep well are more satisfied with life (Shin & Kim, 2018). Income contributes to life satisfaction because it allows someone to meet their physiological needs and work on fulfilling their more advanced needs as people (Masuda et al., 2021). Overall, life satisfaction is an important measure with many contributors that vary by the groups of people being studied. In this paper, we focus on Canadian-born non-immigrants recorded in the 2018 Canadian Community Health Survey (CCHS) survey (Statistics Canada, 2020). The research question this paper seeks to answer is as follows: Does a

higher sense of perceived belonging to local community contribute to higher self-reports of life satisfaction for non-immigrant Canadian-born people in the 2018 CCHS, after controlling for other variables related to life satisfaction? This paper finds that within this population, a higher perceived sense of belonging to local community is consistently associated with higher self-reports of life satisfaction.

II. Data Description

The data set I will be using for this project is the Canadian Community Health Survey (CCHS) 2018 published under the Statistics Canada Open Data License (Statistics Canada, 2020), which is a cross-sectional survey that collects information related to health status, healthcare utilization, and health determinants for persons aged 12 and over living in private dwellings in over 100 health regions covering all provinces and territories (Government of Canada, 2019). For the purposes of my research question, I will be reducing the size of the data to only include Canadian-born non-immigrants, which constitutes my analysis sample, where it is encoded by the variable "sdcdvimm". An important feature of the structure of the data is that ranges of numerical values (for example, the variable "incdgper" meaning "Personal income") are encoded as ordinal numeric values and not the real quantitative values associated with these ranges. In the case of the variable, incdgper, the category "No income or income loss" is recorded as "1", "Less than $20,000" as "2", "$20,000 to $39,999" as "3" and so forth. I use the midpoint method for each of these ranges in order to convert these ordinal observations into quantitative variables for use as one of our regression control variables.

The key variables we will be focusing on are (1) "GEN_010" meaning "Satisfaction with life in general", (2) "GEN_030" meaning "Sense of belonging

to local community", (3) "SLPG005" meaning "Number of hours per night usually spent sleeping", (4) "DHH_SEX" meaning "sex", and (5) "incdgper" meaning "personal income - all sources". "Satisfaction with life in general" is relevant because it is the outcome variable we will run the regression on. This numeric value is the respondent's answer to the question: "Using a scale of 0 to 10, where 0 means "very dissatisfied" and 10 means "Very satisfied", how do you feel about your life as a whole right now?", and provides a gauge of how our survey respondents' quality of life. "sense of belonging to local community" or "perceived_belonging" is the primary variable whose association we are trying to see with life satisfaction, and it is a categorical variable which is encoded in the data as follows: "1" meaning "Very strong", "2" meaning "Somewhat strong", "3": meaning "Somewhat weak", and "4" meaning "Very weak". The "sex" variable is encoded as "1" meaning "male" and "2" meaning "female". These will be converted to factor or dummy variables for the purpose of our regression analysis. The other variables that were included as controls include "Number of hours per night usually spent sleeping", the "sex" variable, and "personal income - all sources". Provided below are summary statistics of our key variables.

Figure 2A: Summary Statistics

```
==============================================================
Statistic                    N       Mean    St. Dev.   Min    Max
--------------------------------------------------------------
general_life_satisfaction   40,989   8.027     1.645      0     10
perceived_belonging         40,989   2.256     0.820      1      4
avg_hours_slept_categories  40,989   7.179     1.378    3.000  12.000
sex                         40,989   1.536     0.499      1      2
personal_income             40,989  41,600.180 27,379.690  0   90,000
--------------------------------------------------------------
```

An important choice that was made in preparation for the regression analysis is that we have chosen to omit "Somewhat weak" as the base level for the levels of perceived belonging because it makes more sense to assume someone has "Somewhat weak" belonging instead of assuming someone has the more extreme

"Very weak" belonging classification. For the "sex" variable we have chosen to omit "male" as the base level. These choices will be important for the interpretation of our regression models.

III. Model

The key assumption for a linear regression model that we need to take is that the true regression equation is linear in its coefficients. The linear regressions we include are (1) a simple regression model, (2) a multiple regression model, (3) a multiple regression model with an interaction term. Our simple regression model is as follows:

(1) $$Y_i = \beta_0 + \beta_1 D_i + \epsilon_i$$

where D_i is our "belonging" variable, and the Y_i term represents the score on the survey question about life satisfaction in general. The β_1 coefficient term captures the effect of perceived belonging on average life satisfaction by taking a difference of means between every other level of perceived belonging: "Very weak", "Somewhat strong", "Very strong" and the base level, "Somewhat weak" perceived belonging to local community. Our multiple regression model is as follows:

(2) $$Y_i = \beta_0 + \beta_1 D_i + \beta_2 G_i + \beta_3 S_i + \beta_4 I_i + \epsilon_i$$

where D_i is our "belonging" variable, G_i is our "sex" variable, S_i is our "belonging" variable, I_i is our "personal income" variable and the Y_i term represents the score on the survey question about life satisfaction in general. The β_1 coefficient term captures the effect of perceived belonging on average life satisfaction after controlling for the effects of average hours of sleep, sex, and income. Our multiple regression model with interaction is as follows:

(3) $$Y_i = \beta_0 + \beta_1 D_i + \beta_2 G_i + \beta_3 D_i \times G_i + \epsilon_i$$

where D_i is our "belonging" variable, G_i is our "sex" variable, and the Y_i term represents the score on the survey question about life satisfaction in general. In this model, the β_1 coefficient term captures the effect of perceived belonging on average life satisfaction controlling for sex and the interaction of the given level of belonging and sex, and the β_3 term captures the impact of having both each category of belonging: "Very weak", "Somewhat strong", "Very strong" and having female status, relative to having male status with "Somewhat weak" belonging. Thus, it allows us to better understand the influence of each level of belonging by seeing if these results also differ by sex.

IV. Table of Results

```
Comparison of Controls
=================================================================
                                         Dependent variable:
                                     ----------------------------
                                     General Life Satisfaction
                                        (1)       (2)       (3)
-----------------------------------------------------------------
Very Weak Belonging                  -0.674*** -0.610*** -0.719***
                                     (0.032)   (0.032)   (0.046)

Somewhat Strong Belonging             0.470***  0.457***  0.454***
                                     (0.019)   (0.019)   (0.028)

Very Strong Belonging                 0.825***  0.821***  0.786***
                                     (0.025)   (0.024)   (0.037)

Female                                          0.113***  0.018
                                               (0.016)   (0.031)

Very Weak Belonging and Female                            0.092
                                                         (0.064)

Somewhat Strong Belonging and Female                      0.029
                                                         (0.038)

Very Strong Belonging and Female                          0.069
                                                         (0.050)

Constant                              7.712***  6.298***  7.702***
                                     (0.015)   (0.046)   (0.022)

-----------------------------------------------------------------
Avg. Hours of Sleep Controls            No        Yes       No
Personal Income Controls                No        Yes       No
Observations                          40,989    40,989    40,989
R2                                     0.058     0.086     0.058
Adjusted R2                            0.058     0.086     0.058
=================================================================
Note:                                   *p<0.1; **p<0.05; ***p<0.01
```

V. Discussion

From our simple regression model whose results are displayed in column (1), we can see that average life satisfaction for respondents with "Somewhat weak" belonging to local community is a score of 7.712 on a scale of 0 to 10. This simple regression model suggests that having a "Very weak" sense of belonging to local community, compared to the base level of "Somewhat weak" belonging, is associated with a -0.674 decrease in general life satisfaction score. A "Somewhat strong" sense of belonging to local community, compared to the base level of "Somewhat strong" belonging, is associated with a 0.47 increase in general life satisfaction score. A "Very strong" sense of belonging to local community, compared to the base level of "Somewhat weak" belonging, is associated with a 0.825 increase in general life satisfaction score. All of these above are reported with p-values of 0.01, providing strong evidence for these estimates. The r-squared and adjusted r-squared value is low at 0.058, which means that only a small proportion of the variance in general life satisfaction is explained by the sense of belonging to local community in this simple regression model.

From our multiple regression model whose results are displayed in column (2), we can see that average life satisfaction for respondents with "Somewhat weak" belonging to local community, who are "male", have 3 average hours of sleep, and personal income of $0, have a score of 6.298 on a scale of 0 to 10 for general life satisfaction. The table suggests that having a "Very weak" sense of belonging to local community, compared to the base level of "Somewhat weak" belonging and other aforementioned base characteristics, is associated with a -0.61 decrease in general life satisfaction score, controlling for sex, average hours of sleep, and personal income. A "Somewhat strong" sense of belonging to local community, compared to the base level of "Somewhat weak" belonging and other aforementioned base characteristics, is associated with a 0.457 increase in general

life satisfaction score, controlling for sex, average hours of sleep, and personal income. A "Very strong" sense of belonging to local community, compared to the base level of "Somewhat weak" belonging and other aforementioned base characteristics, is associated with a 0.821 increase in general life satisfaction score, controlling for sex, average hours of sleep, and personal income. All of these above are reported with p-values of 0.01, providing strong evidence for these estimates, and showing their statistical significance. The r-squared and adjusted r-squared value is low at 0.086, which means that only a small proportion, 8.6%, of the variance in general life satisfaction is explained by the perceived belonging, sex, average hours of sleep, and personal income in this model.

Finally, from our multiple regression model with interaction whose results are displayed in column (3), we can see that average life satisfaction for respondents with "Somewhat weak" belonging to local community, who are male have a score of 7.702 on a scale of 0 to 10 for general life satisfaction. The table suggests that being a male with a "Very weak" sense of belonging to local community, compared to the base level of a male with "Somewhat weak" belonging, is associated with a -0.719 decrease in general life satisfaction score. Being a male with a "Somewhat strong" sense of belonging to local community, compared to the base level of a male with "Somewhat weak" belonging, is associated with a 0.454 increase in general life satisfaction score. Further, being a male with a "Very strong" sense of belonging to local community, compared to the base level of a male with "Somewhat weak" belonging, is associated with a 0.786 increase in general life satisfaction score. The table also shows the coefficient on the interaction term, which tells us that having both "Very weak" belonging and female status is associated with a decrease in life satisfaction score by -0.719+0.092= -0.627 relative to a female with "Somewhat weak" belonging.

Having both "Somewhat strong" belonging and female status is associated with an increase in life satisfaction score by 0.454+0.029 = 0.483 relative to a female with "Somewhat weak" belonging. Lastly, having both "Very strong" belonging and female status is associated with an increase in life satisfaction score by 0.786+0.069 = 0.855 relative to being a female with "somewhat weak" belonging. All of these above are reported with p-values of 0.01, providing strong evidence for these estimates. Ultimately, we see that in each of these models, higher levels of belonging are associated with higher levels of life satisfaction. A drawback of our analysis is that while we can say higher perceived belonging is associated with higher life satisfaction, we cannot say it is causal: that higher perceived belonging causes higher life satisfaction.

VI. Specification Check

One issue with our current regression specification is that the adjusted r-squared value is quite low, at 0.086, meaning sex and perceived belonging only explain 8.6% of the variation in general life satisfaction. Thus, a logical next step would be to try adding more control variables that are relevant in explaining Y_i. Perceived health is found to be a predictor of life satisfaction (Palmore & Luikart, 1972), and life stress has been found to be a moderate predictor of life satisfaction (Hamarat et al., 2001). In the Figure 6A below, we find that adding "perceived health" and "perceived life stress" as controls in our multiple regression model increases our adjusted r-squared from 0.086 in the original multiple regression model to a significantly higher 0.333, meaning 33.3% of the variation in general life satisfaction is explained by our updated model with more controls.

Figure 6A: Specification Check

```
Multiple Regression
========================================================
                                Dependent variable:
                            ----------------------------
                            General Life Satisfaction
--------------------------------------------------------
Very Weak Belonging                   -0.352***
                                       (0.027)

Somewhat Strong Belonging              0.309***
                                       (0.016)

Very Strong Belonging                  0.592***
                                       (0.021)

Constant                               8.408***
                                       (0.049)


--------------------------------------------------------
Sex Controls                             Yes
Avg. Hours of Sleep Controls             Yes
Personal Income Controls                 Yes
Perceived Health Controls                Yes
Perceived Life Stress Controls           Yes
Observations                           40,882
R2                                      0.333
Adjusted R2                             0.333
========================================================
Note:                         *p<0.1; **p<0.05; ***p<0.01
```

VII. Testing for Heteroskedasticity

Figure 7A: Testing Heteroskedasticity with the Breusch-Pagan Test

```
Call:
lm(formula = resid_sq ~ vweak_belonging_dummy + somestrong_belonging_dummy +
    verystrong_belonging_dummy + female_dummy + avg_hours_slept_categories +
    personal_income + perceived_health + perceived_life_stress,
    data = CCHS_clean_new_sample)

Residuals:
<Labelled double>: Satisfaction with life in general
   Min    1Q Median    3Q    Max
-8.512 -1.361 -0.790  0.308 95.770

Labels:
 value      label
    96 Valid skip
    97 Don't know
    98    Refusal
    99 Not stated

Coefficients:
                                        Estimate Std. Error t value Pr(>|t|)
(Intercept)                              1.98615    0.13564  14.643  < 2e-16 ***
vweak_belonging_dummy                    1.00655    0.07568  13.300  < 2e-16 ***
somestrong_belonging_dummy              -0.15586    0.04438  -3.512 0.000445 ***
verystrong_belonging_dummy              -0.04090    0.05832  -0.701 0.483150
female_dummy                            -0.06676    0.03781  -1.766 0.077419 .
avg_hours_slept_categories              -0.02486    0.01356  -1.833 0.066861 .
personal_income                         -0.12514    0.01403  -8.916  < 2e-16 ***
perceived_healthVery good                0.07160    0.05014   1.428 0.153317
perceived_healthGood                     0.48672    0.05285   9.210  < 2e-16 ***
perceived_healthFair                     1.69395    0.07302  23.200  < 2e-16 ***
perceived_healthPoor                     4.32439    0.11282  38.329  < 2e-16 ***
perceived_life_stressNot very stressful -0.21152    0.06000  -3.526 0.000423 ***
perceived_life_stressA bit stressful    -0.03607    0.05633  -0.640 0.521968
perceived_life_stressQuite a bit stressful 0.20202  0.06553   3.083 0.002050 **
perceived_life_stressExtremely stressful 1.68079    0.12165  13.817  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.712 on 40867 degrees of freedom
Multiple R-squared:  0.07969,   Adjusted R-squared:  0.07938
F-statistic: 252.8 on 14 and 40867 DF,  p-value: < 2.2e-16
```

Inspecting the results of the Breusch-Pagan Test, we can see from the *F*-statistic that we can strongly reject the assumption of homoskedasticity since our results are denoted by 3 asterisks. Since this data is heteroskedastic because the residuals can be predicted using the explanatory variables, we need to correct for heteroskedasticity with robust standard errors.

VIII. Correcting Heteroskedasticity with Robust Standard Errors

In order to correctly estimate my standard errors, we will correct our standard errors with robust standard errors. Provided below is a comparison of the standard model and the model with robust standard errors.

Figure 8A: Comparison of Non-Robust & Robust Standard Errors Model

```
=============================================================
                                   Dependent variable:
                             --------------------------------
                             General Life Satisfaction
                                DEFAULT         ROBUST
                                  (1)             (2)
-------------------------------------------------------------
Very Weak Belonging            -0.352***       -0.352***
                                (0.027)         (0.035)

Somewhat Strong Belonging       0.309***        0.309***
                                (0.016)         (0.016)

Very Strong Belonging           0.592***        0.592***
                                (0.021)         (0.021)

Constant                        8.408***        8.408***
                                (0.049)         (0.055)

-------------------------------------------------------------
Sex Controls                      Yes             Yes
Avg. Hours of Sleep Controls      Yes             Yes
Personal Income Controls          Yes             Yes
Perceived Health Controls         Yes             Yes
Perceived Life Stress Controls    Yes             Yes
Observations                    40,882          40,882
R2                               0.333           0.333
Adjusted R2                      0.333           0.333
=============================================================
Note:                       *p<0.1; **p<0.05; ***p<0.01
```

These corrections make our standard errors more accurate. All of the estimates are the same between the non-robust and robust model, except that the corrected very weak belonging level has an increased standard error, from 0.027 in the

original model to 0.035 in the robust model, and the constant term has an increased standard error from 0.049 to 0.055.

IX. Evaluating Robustness by Changing the Outcome Variable

In order to evaluate whether or not the results we have generated answer the initial research question, we can perform the same regression on another measure of life satisfaction in the CCHS, "gendvswl" described as "Satisfaction with life in general" with observations ranging from "Very Satisfied", "Satisfied, "Neither satisfied nor dissatisfied", "Dissatisfied", to "Very Dissatisfied". We convert these categories into numeric observations out of 10, with "Very Satisfied" transformed to 10, "Satisfied" to 7.5, "Neither satisfied nor dissatisfied" to 5, "Dissatisfied" to 2.5, and "Very Dissatisfied" transformed to 0, in order to run a regression with this variable as the outcome variable.

The table on the left displays the initial simple regression, and the table on the right displays this unique robustness check regression. We find that the results of the initial simple regression are consistent with this other measure of general life satisfaction, as depicted below, with minor numerical variations. This can be seen with the "Very weak" belonging being correlated with lower life satisfaction compared to "Somewhat weak" belonging, "Somewhat strong" and "Very strong" belonging correlated with higher life satisfaction compared to "Somewhat weak" belonging. The magnitudes of these correlations are also similar, and both are statistically significant, despite the table on the outcome variable for the table on the right being generated from a transformation of a categorical variable. This provides further evidence for the robustness of the provided estimates in the initial table.

Figure 9A: Initial Multiple Regression		
==		
	Dependent variable:	

	General Life Satisfaction	
	DEFAULT	ROBUST
	(1)	(2)
--		
Very Weak Belonging	-0.352***	-0.352***
	(0.027)	(0.035)
Somewhat Strong Belonging	0.309***	0.309***
	(0.016)	(0.016)
Very Strong Belonging	0.592***	0.592***
	(0.021)	(0.021)
Constant	8.408***	8.408***
	(0.049)	(0.055)
--		
Sex Controls	Yes	Yes
Avg. Hours of Sleep Controls	Yes	Yes
Personal Income Controls	Yes	Yes
Perceived Health Controls	Yes	Yes
Perceived Life Stress Controls	Yes	Yes
Observations	40,882	40,882
R2	0.333	0.333
Adjusted R2	0.333	0.333
==		
Note:	*p<0.1; **p<0.05; ***p<0.01	

Figure 9B: Unique Multiple Regression		
==		
	Dependent variable:	

	General Life Satisfaction	
	DEFAULT	ROBUST
	(1)	(2)
--		
Very Weak Belonging	-0.373***	-0.373***
	(0.031)	(0.039)
Somewhat Strong Belonging	0.308***	0.308***
	(0.018)	(0.018)
Very Strong Belonging	0.591***	0.591***
	(0.024)	(0.024)
Constant	8.634***	8.634***
	(0.053)	(0.058)
--		
Sex Controls	Yes	Yes
Avg. Hours of Sleep Controls	Yes	Yes
Personal Income Controls	Yes	Yes
Perceived Health Controls	Yes	Yes
Perceived Life Stress Controls	Yes	Yes
Observations	40,882	40,882
R2	0.292	0.292
Adjusted R2	0.291	0.291
==		
Note:	*p<0.1; **p<0.05; ***p<0.01	

X. Conclusion

In conclusion, we find that while perceived belonging to local community contributes to life satisfaction in general after controlling for sex, average hours of sleep, personal income, perceived health, and perceived life stress, it is only one of many factors that contribute to general life satisfaction: it explains only 5.8% of the variation in general life satisfaction. Conversely, the combination of perceived belonging, sex, average hours of sleep, personal income, perceived health, and perceived life stress together explain 33.3% of the variation in general life satisfaction for this population. This result is substantiated by the fact that the value of both r-squared and adjusted r-squared are equal, meaning r-squared did not go up only because we added additional control variables.

Ultimately, we consistently find that having stronger levels of belonging relative to the somewhat weak belonging level is associated with increases in general life satisfaction, and having very weak belonging relative to having somewhat weak

belonging is associated with lower general life satisfaction in both the simple and controlled regression models through our study of Canadian-born non-immigrants. This result is robust across different measures of general life satisfaction provided in the CCHS.

VIII. References

Government of Canada, S. C. (2019, October 22). Canadian Community Health
  Survey: Combined Data, 2017 and 2018. Canadian Community Health
  Survey: Combined data, 2017 and 2018. Retrieved March 5, 2023, from
  https://www150.statcan.gc.ca/n1/daily-quotidien/191022/dq191022d-eng.htm

Hamarat, Dennis Thompson, Karen M. Zabrucky, Don Steele, Kenneth B.
  Matheny, Ferda Aysan, Errol. (2001). Perceived stress and coping resource
  availability as predictors of life satisfaction in young, middle-aged, and older
  adults. Experimental Aging Research, 27(2), 181-196.
  https://doi.org/10.1080/036107301750074051'

Masuda, Y. J., Williams, J. R., & Tallis, H. (2021). Does life satisfaction vary with
  time and income? Investigating the relationship among free time, income, and
  life satisfaction. Journal of happiness studies, 22, 2051-2073.

Palmore, E., & Luikart, C. (1972). Health and social factors related to life
  satisfaction. Journal of Health and Social Behavior, 13(1), 68-80.
  https://doi.org/10.2307/2136974

Shin, J., & Kim, J. K. (2018). How a good sleep predicts life satisfaction: The role
  of zero-sum beliefs about happiness. Frontiers in Psychology, 9, 1589-1589.
  https://doi.org/10.3389/fpsyg.2018.01589

Solan, M. (2017, October 5). The secret to happiness? here's some advice from
  the longest-running study on Happiness. Harvard Health. Retrieved April 10,
  2023, from

https://www.health.harvard.edu/blog/the-secret-to-happiness-heres-some-advice-from-the-longest-running-study-on-happiness-2017100512543

Statistics Canada. (2020). Canadian Community Health Survey: Public Use Microdata File, 2017/2018 - ARCHIVED, June 2020. [Data file]. Retrieved from https://www150.statcan.gc.ca/n1/en/catalogue/82M0013X2020001

Veenhoven, R. (1996). Developments in satisfaction-research. Social Indicators Research, 37(1), 1-46. https://doi.org/10.1007/BF00300268