School Ranking and Student Commute as a Factor in Post-Secondary School Choice

by

Laura Cline, Ailin Z.W. Li, Lily (Lillian) Shaddick, Will Trefiak, Giselle Wenban

A final paper submitted in conformity with the requirements

for the course of INF1344H

Faculty of Information

University of Toronto

Abstract

This paper aims to examine the relationship between commute time as a factor in students' postsecondary school selection and the ranking of the postsecondary institution they chose to attend, using data from the StudentMoveTO's Fall 2019 dataset. Our hypothesis is that students' self-reported perception that commute time impacted their school choice will be associated with attending a lower-ranked postsecondary school. Linear regression analysis was performed to determine if a student's postsecondary school ranking was dependent on if their school choice was influenced by commute time, while controlling for network commute distance, gender, age, living situation, and family income level. Based on the results of this analysis, students that chose their postsecondary school based on commute time were more likely to attend a lower ranking school. However, our results found that family income level was also a significant predicting variable. Additionally, we found that students who were older men living with a host family or at a friend's house were more likely to attend a lower ranking school.

Table of Contents

List of Tables

List of Figures

1. Introduction

There are a variety of barriers that students experience and considerations they must make when choosing a postsecondary institution (whether college or university). Many students opt to live at home in order to save money, many work part-time jobs to support themselves through school, etc. As will be explored in the literature review, students may sacrifice going to a more prestigious postsecondary and choose either a college or a lower-ranked university closer to home, for the sake of saving money on tuition and living expenses. Commute time and cost are another factor in this consideration; a student may choose a school that is closer to their home, disregarding that school's reputation, for the sake of saving time commuting to a better but further away school. Thus, this paper investigates whether there is a measurable difference in the rank of postsecondary school attended depending on whether students face commuting barriers.

1.1 Research Question and Hypotheses

1.1.1 Research Question

Within the Greater Toronto and Hamilton Area (GTHA), is the self-perceived impact of commuting on school choice associated with the rank of post-secondary institution (college and university) that students attend, controlling for network commute distance, age, gender, living situation, and annual family income?

1.1.2 Hypotheses

H_0: Students' self-reported perception that their commute influenced their school choice is not associated with the overall ranking of the postsecondary institution these students attend.

H_1: A student's self-reported perception that commute does impact school choice will be associated with attending a lower-ranked post-secondary school.

2. Theoretical Framework & Existing Literature

The literature that informs this report includes Briggs (2006), whose study used data collected from 651 Scottish undergraduate students to analyze the factors that influenced their choice of postsecondary school. The study found that the top three factors that influenced students' school choice included school reputation, distance from home and school location.

The second study is Frenette (2004), who collected data from high school students to study the factors that influenced their decision to attend either a college or a university. The study also sought to answer whether commute distance posed a barrier to attending postsecondary education. It found that students are less likely to attend a college or university if they are located beyond 80km from the campus, with students from lower income families being disproportionately affected.

Third, the NYC Urban Institute (2018) studied the influence of postsecondary students' average commute time and the choice of postsecondary institution they attended within New York City. The study found that, on average, students chose to attend postsecondary institutions that were within a 35.5 minute commute from their home. The study also analyzed the commute times for all of the postsecondary institutions these students applied to. They found that the average travel time for each students' top three college applications was 36.4 minutes. The authors argue that students who had the longest commute times to NYC's most prestigious schools (e.g., Columbia University and New York University) were the least likely to attend these schools. Instead, these students were more likely to attend less prestigious public and private colleges in the same borough. Thus, the authors concluded that commute times affect students' postsecondary school choice.

3. Methods

3.1 Data Set

We used the StudentMoveTO's dataset on their survey about student travel patterns and commutes from Fall 2019. The survey was developed by the University of Toronto's Transportation Research Institute (UTTRI) and the School of Cities. The survey questionnaire collected information on personal and household attributes, institutional information, mobility tools used for daily travel, a complete travel diary of a day, attitudes and behaviours towards key transportation issues and policies in the region, as well as perceptions of travel satisfaction, campus participation, academic success, and subjective well-being. (Link to dataset: http://www.studentmoveto.ca/resources-2/2019-survey/.) The data has already been cleaned by the researchers.

In the GTHA, there are over 600,000 postsecondary students who attend one of the 10 postsecondary institutions represented in this dataset. Our sample size will be 10,631 of the survey respondents, who are postsecondary students that live in the GTHA and attend one of 10 universities and colleges in the GTHA, including: Centennial College, Durham College, McMaster University, OCAD University, Ontario Tech University, Ryerson University, Sheridan College, Mohawk College, University of Toronto, and York University. The data was collected between October 1st to November 30th, 2019.

Within this dataset, students had a network commute distance between 1.44 meters and 395,474.53 meters (in other words, 395.47 kilometers). For the statement "My commute was a factor in deciding which school to attend", 5,409 students answered "Yes" and 5,222 students answered "No". The other 8,504 students did not answer the question.

We also used the Webometrics Ranking of World Universities (2020) ranking data, as it aggregates rankings for both colleges and universities (versus only colleges or only universities). This ranking system mainly weighs academic contributions, including number of open source publications and impact factor of publications.

3.2 Measurement

3.2.1 Independent Variables

The independent variable is *commute influencing school choice*, which will be measured based on “ps05commuteFactorInSchoolChoice” in the StudentMoveTO data set. Students reported whether they believed commute time was a factor for them when selecting a postsecondary institution. Students were asked if they agree with the following statement: “My commute was a factor in deciding which school to attend.” For this analysis, responses are labeled as “No” = 1 and “Yes” = 2.

3.2.2 Dependent Variables

The dependent variable is *college/university ranking*, which will be measured based on “psinstitution” in the StudentMoveTO data set. In the data set, a student’s postsecondary school is listed based on the institution that sent the student the survey invitation. The institution options are the ten postsecondary schools in the GTHA. For the present research purposes, a level will be assigned to each institution based on the Webometrics Ranking of World Universities (2020) data, with higher ranking institutions receiving a lower magnitude. The rankings are as follows: “University of Toronto” = 1, “McMaster University” = 2, “York University” = 3, “Ryerson University” = 4, “Sheridan College Institute of Technology and Advanced Learning” = 5, “OCAD University” = 6, “Centennial College Toronto” = 7, “Mohawk College” = 8, “Ontario Tech University” = 9, and “Durham College” = 10. In this manner, the categorical variable is revised to become an ordinal variable once levels are introduced.

3.2.3 Control Variables

The first control variable is network *commute distance*. Commute distance will be measured based on “networkCommuteDistance” in the dataset, which indicates the distance (in meters) between a student’s home and their postsecondary school’s main campus based students’ self-reported street-route to school.

The second control variable is *gender*, measured by “psgender,” with possible responses being “Female”; “Male”; “Non-Binary / third gender”; and “prefer not to answer”. The response “prefer not to answer” will be filtered out of valid responses before analysis is conducted.

The third control variable is *age*, measured by “psage,” with possible responses being a number between 1-200.

The fourth control variable is *household living situation*, measured by “hhlivingsituation,” with possible responses being “Live with family/parents”; “Live with partner”; “Live alone”; “Live with roommates”; and “Live with host family or at a friend’s house”.

The last control variable is *family income level*, measured by "familyincomelevel" in the dataset, with possible responses being "Less than $ 14,999"; "$ 15,000 - 29,999"; "$ 30,000 - 39,999"; "$ 40,000 - 49,999"; "$ 50,000 - 59,999"; "$ 60,000 - 69,999"; "$ 70,000 - 79,999"; "$ 80,000 - 89,999"; "$ 90,000 - 99,999"; "$ 100,000 - 124,999"; "$ 125,000 - 149,999"; "$ 150,000 - 199,999"; "$ 200,000 +"; "I don't know"; and "Prefer not to answer". Before analysis is conducted, responses "I don't know" and "Prefer not to answer" will be filtered out.

3.3 Data Analysis

Descriptive statistics were initially utilized to explore linearity between our dependent and independent variables using tableau. After establishing these exploratory measures, our team conducted an OLS regression to determine collinearity between all our dependent, independent, and control variables. Given the multiple categories of our dependent variable, we determined a linear regression is the optimal analytical approach for determining potential correlations between our variables compared to a logistic regression, which requires a nominal dependent variable. In addition, the linear relationships observed in our exploratory analysis also played a factor in determining the approach to select. Dummy variables were created for *psgender*, *psage*, *hhlivingsituation*, and *familyincomelevel* and the reference level for the latter two variables were also adjusted to correspond with the modes of our cleaned sample. As mentioned above, we also recoded our independent variable and set University of Toronto as the reference level. Data cleaning and statistical analysis was conducted using the R programming language in RStudio.

4. Results & Discussion

4.1 Descriptive Results

Descriptive statistical analysis was conducted to better understand and meaningfully summarize the data, and to highlight interesting findings. As shown in Tables 1, 2, and 3 below, all variables have a sample size of 5358.

As seen in Table 1 below, the mode of dependent variable College/University Ranking was #1, which indicates that the most common institution attended by students (n = 5358) was the highest ranked postsecondary school in the GTHA, the University of Toronto.

Table 1. Dependent Variable

Dependent Variable	**N**	**Descriptive Statistic**
College/University Ranking (psinstitution)	5358	Mode = #1

For the independent variable *commute influencing school choice*, students (n = 5358) most frequently responded "Yes" meaning that commute did in fact influence school choice for the majority. Figure 2. Below shows the breakdown of responses (either "Yes" or "No") regarding commute influencing school choice for each institution in the GTHA.

Table 2. Independent Variable

Independent Variable	**N**	**Descriptive Statistic**
Commute influencing school choice (ps05commuteFactorInSchoolChoice)	5358	Mode: Yes

Table 3 below highlights important descriptive statistics for the five control variables: commute distance, gender, age, household living situation, and family income level. For control variable *commute distance*, students' (n = 5358) mean commute distance was 17699.63 meters (SD = 19991.44 metres). While the average commute to school is approximately 17.69km, the high standard deviation indicates that there are observations very dispersed from the mean, which will be elaborated on further below in the discussion of Figure 1.
Furthermore, for students (n = 5358) in our data set, most respondents were female, the mean age was 21.59 years (SD = 5.36), the most frequent living situation was with family/parents, and the most frequent household income was $15,000 – $29,999.

Table 3. Control Variables

Control Variables	**N**	**Descriptive Statistic**
Commute distance (networkCommuteDistance)	5358	Mean: 17699.63 meters SD: 19991.44 meters
Gender (psgender)	5358	Mode: Female
Age (psage)	5358	Mean: 21.5991 years SD: 5.362688 years
Household living situation (hhlivingsituation)	5358	Mode: Live with Family/Parents
Family income level (familyincomelevel)	5358	Mode: $15,000 – $29,999

Figure 1 below shows a boxplot for the control variable Network Commute Distance (networkcommutedistance). The box plot shows that the median network commute distance is

12,383 metres. The lower quartile is 2,932 meters and the upper quartile is 25,878 meters. The box plot indicates a minimum (Q1 - 1.5*IQR) of 1 meter and a maximum (Q3 + 1.5*IQR) of 60,167 meters. There are multiple outliers, and the most extreme outlier (i.e longest network commute distance in the data set) is 395,748 metres.

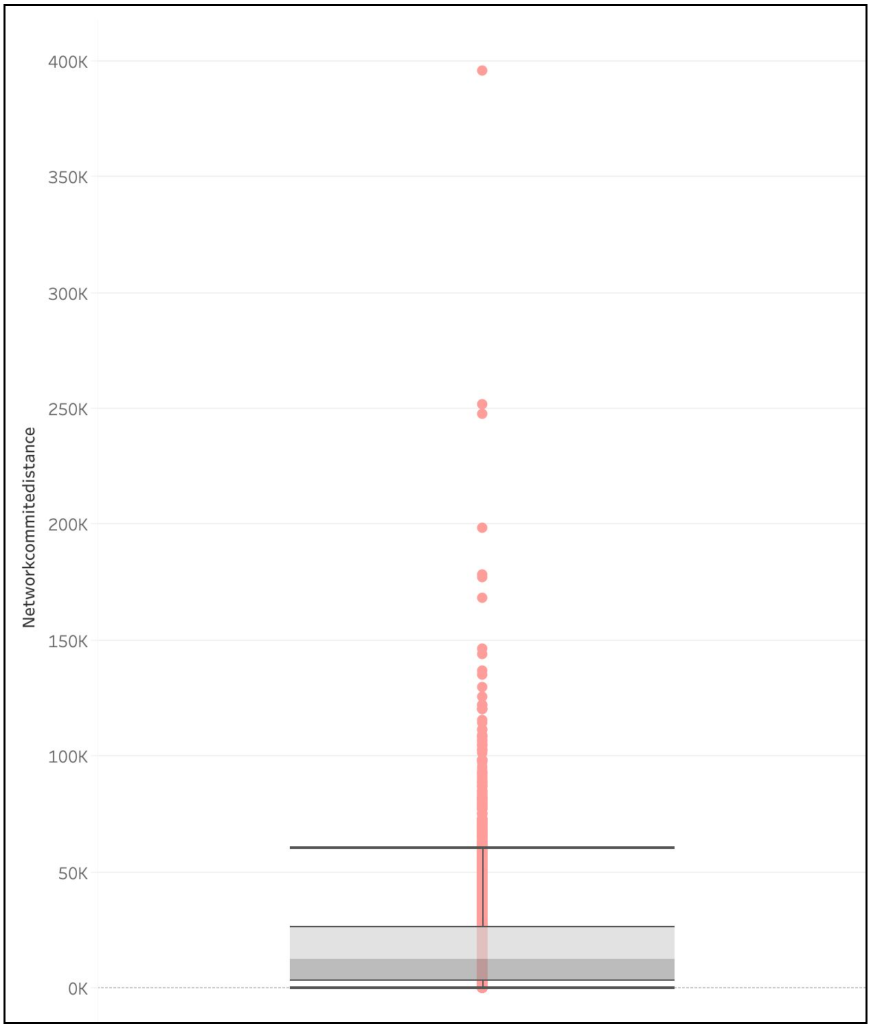

Figure 1. Boxplot for the control variable Network Commute Distance (networkcommutedistance).

The boxplot shows that the median is closer to the bottom of the box, and the lower whisker is very short. These features indicate that the boxplot distribution is positively skewed (skewed right). The Mean for network commute distance is 17699.63 metres, which is greater than the Median 12,383 metres shown on the box plot. The boxplot also shows the extreme outliers that contribute to the high Standard Deviation for network commute distance (SD = 19991.44).

The bar graph in Figure 2 below shows the breakdown of “Yes” or “No” responses regarding the question “Commute Time Affected School Choice” for each institution in the GTHA. Here is the link to some additional graphs to explore the data in Tableau:

https://public.tableau.com/profile/laura.cline?fbclid=IwAR1dN4dYREg9VVD8-w3hhXGyxgvCJjTCfdhJ6fTrt8DN2aUbTLJ9TcBtYkQ#!/vizhome/StudentMoveTOCommutesAffectonSchoolChoice/Story1?publish=yes.

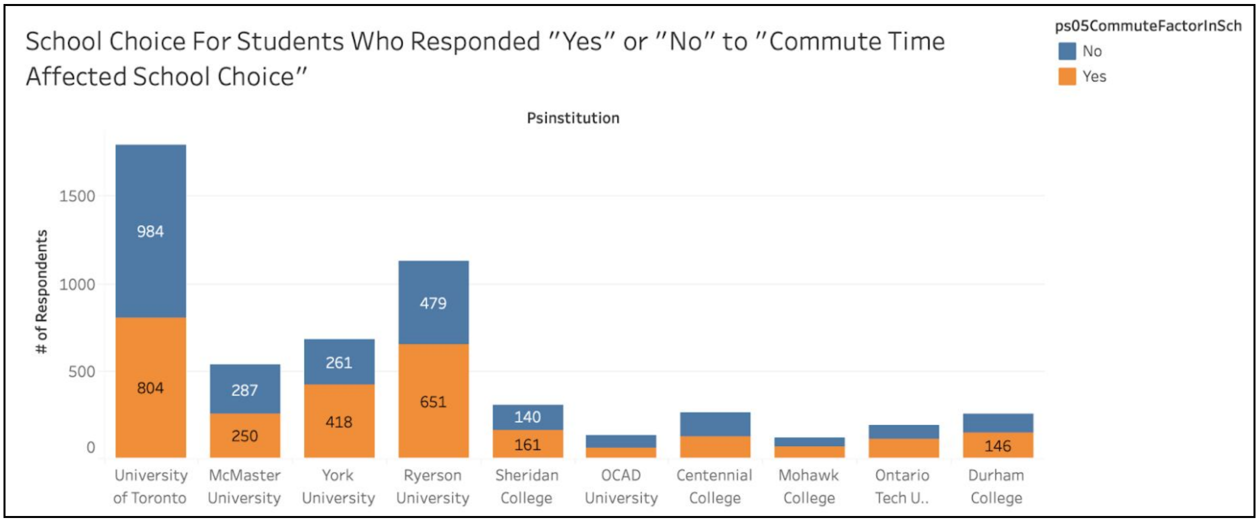

Figure 2. Bar graph of responses either "Yes" or "No" regarding "Commute Time Affected School Choice" (ps05commuteFactorInSchoolChoice) for each postsecondary institution (psinstitution). View interactive graph.

4.2 Linear Regression Results

A single linear regression model (Model 1) was built to examine the relationship between our dependent and independent variable, and each of the control variables. Based on the R2 value of Model 1, it explains about 6.08% of the variation in post-secondary school ranking. Table 5 displays each of the coefficients obtained for the independent and control variables.

Table 4. Linear Regression Results Summary

Residual Standard Error	2.514 on 5336 degrees of freedom
Multiple R Squared	0.06083
Adjusted R Squared	0.05713
F-Statistic	16.46 on 21 and 5336 degrees of freedom
p-value	0.000
Residual Standard Error	2.514 on 5336 degrees of freedom

Table 5. Linear Regression Results per Variable

Model		B	Standard Error	t	Significance
1	(Constant)	2.939	0.196	15.017	**0.000 *****
	Commute Factor in School Choice.f *Yes*	0.307	0.074	4.122	**0.000 *****
	Network Commute Distance	0.000	0.000	2.019	**0.044 ***
	Gender.f *Male*	0.310	0.074	4.188	**0.000 *****
	Gender.f *Binary/Third Gender*	0.612	0.302	2.028	**0.043 ***
	Age	0.035	0.007	5.029	**0.000 *****
	Living Situation.f *Live Alone*	-0.806	0.147	-5.469	**0.000 *****
	Living Situation.f *Live With Host Family or at Friend's House*	1.182	0.265	4.461	**0.000 *****
	Living Situation.f *Live With Partner*	-0.117	0.160	-0.734	0.463
	Living Situation.f *Live With Roommates*	-0.151	0.096	-1.581	0.113
	Familyincomelevel.f *Less than $14,999*	0.578	0.156	3.700	**0.000 *****
	Familyincomelevel.f *$30,000 – $39,000*	-0.074	0.159	-0.468	0.640
	Familyincomelevel.f *$40,000 – $49,999*	-0.459	0.157	-2.918	**0.004 ****

	Familyincomelevel.f *\$50,000 – \$59,999*	-0.730	0.154	-4.749	**0.000 *****
	Familyincomelevel.f *\$60,000 – \$69,999*	-0.503	0.160	-3.151	**0.001 ****
	Familyincomelevel.f *\$70,000 – \$79,999*	0.527	0.172	-3.070	**0.002 ****
	Familyincomelevel.f *\$80,000 – \$89,999*	-0.669	0.174	-3.844	**0.000 *****
	Familyincomelevel.f *\$90,000 – \$99,999*	-0.918	0.174	-5.205	**0.000 *****
	Familyincomelevel.f *\$100,000 – \$124,999*	-0.716	0.147	-4.861	**0.000 *****
	Familyincomelevel.f *\$125,000 – \$149,999*	-0.856	0.178	-4.804	**0.000 *****
	Familyincomelevel.f *\$150,000 – \$199,999*	-1.027	0.178	-5.784	**0.000 *****
	Familyincomelevel.f *\$200,000 +*	-1.271	0.175	-7.252	**0.000 *****

Significance Codes: $* p < 0.05$; $** p < 0.01$; $*** p < 0.001$

An intercept of 2.939 means that at reference level—in this case, for a young female who has a commute distance of 0m, whose commute did not influence school choice, and is living with her family who has an annual household income of $15,000–29,999—her school rank is 2.939, which is a relatively high-ranking school in our model, since lower ranked schools have a higher value.

The coefficient for our independent variable, commute influencing school choice, is also significant, such that the average school ranking for students whose commute influenced their school choice is 0.307 units higher than for students whose commute did not influence their school choice. In our model, this means a slight decrease in school ranking for students whose commutes influenced their school choice.

All of the control variables in Model 1 also produced significant results. Each of the following results should be interpreted as holding all other variables constant. First, an increase in *network commute distance* was correlated with a decrease in school ranking. Even though this coefficient is statistically significant, this result is not practically meaningful as the coefficient value is 0.000004.

Second, older people are more likely to attend a lower-ranked post-secondary, with 1 unit of age correlated with a 5.029 unit increase on school ranking (which, in our case, is a lower-ranked post-secondary).

Third, compared to women, men are more likely to attend a lower-ranked post-secondary, with a 0.310 unit increase for the variable school ranking. Individuals who identified as a third gender or non-binary also saw a 0.612 unit increase in the school ranking variable when compared to women, again, meaning a lower ranked school.

Fourth, those who live with a host family/at a friend's house are more likely to attend a lower-ranked post-secondary, seeing a 1.182 unit increase on the school ranking variable, compared to students who live at home. Those living alone, though, are more likely to attend a higher-ranked post-secondary, seeing a 0.806 unit decrease on the school ranking variable, compared to students who live at home. The other two living conditions, living with a partner and living with a roommate, did not produce significant results.

And lastly, almost all of the family income variables produced significant results. Compared to students whose families have an annual household income of $15,000–$29,999, those whose family income is less than $14,999/year are more likely to attend a lower-ranked post-secondary, with a 0.578 unit increase on the school ranking variable. On the other hand, those whose family income is above $40,000/year are more likely to attend a higher-ranked post-secondary, with a 0.459 unit decrease on the school ranking variable when compared to students with an annual household income of $15,000–$29,999. The results for each subsequent income bracket are also significant, with results increasing in significance at the $80,000+ range.

5. Conclusion

Based on these results, we reject the null hypothesis, which was: students' self-reported perception that their commute influenced their school choice is not associated with the overall ranking of the postsecondary institution these students attend. Significant predicting variables for choice of post-secondary institution point to family income, wherein students from lower income households are more likely to attend lower ranked institutions. In addition, the perceived impact of commute distance on students' choice of school has a measurable effect on the rank of the post-secondary institutions that students attend. This indicates that similar patterns exist between students of the GTHA and the existing studies on factors influencing post-secondary school choice, as detailed in the literature of Briggs (2006), Frenette (2004) and the NYC Urban Institute (2018). New insights for contribution to the StudentMoveTO study also include the control variables of age, gender and living situation; men, older individuals and individuals living with a host family or at a friend's house are more likely to attend a lower ranked institution.

Using these results, we recommend improving public transportation infrastructure in order to reduce commute times and provide more affordable public transit options for post-secondary students. These changes will increase students' ability to reach all campuses effectively and in a timely manner at all hours of the day so they can attend higher ranking postsecondary institutions in the GTHA, have access to on-campus opportunities and services, and live in more affordable housing options outside the downtown core. This will not only allow Universities and Colleges to attract the best and brightest students, but will also ensure that potential students have an equal opportunity to attend the best programs in the GTHA and achieve their academic goals regardless of income, living situation, family structure, disabilities, and other factors. These changes will also have a positive economic impact by building a more educated, high-skilled and talented workforce. Therefore, universities, colleges, and city planning authorities must collaborate to respond to students' transportation needs and challenges by incorporating these issues into campus planning and urban policy.

References

Briggs, Senga (2006). An exploratory study of the factors influencing undergraduate student choice: the case of higher education in Scotland. *Studies in Higher Education*, Vol. 31, No. 6, December 2006, pp. 705–722. Retrieved October 28, 2020, from https://journals-scholarsportal-info.myaccess.library.utoronto.ca/pdf/03075079/v31i0006/705_aesotfcoheis.xml.

Corcoran, S. P. (2018). School Choice and Commuting: How Far New York City Students Travel to School. *NYC Urban Institute*, 1-58. Retrieved October 28, 2020, from https://www.urban.org/sites/default/files/publication/99205/school_choice_and_commuting_3.pdf.

Frenette, Marc (2004). Access to College and University: Does Distance to School Matter?. *Canadian Public Policy / Analyse de Politiques*, Vol. 30, No. 4 (Dec., 2004), 427-443. Retrieved October 28, 2020, from https://www-jstor-org.myaccess.library.utoronto.ca/stable/3552523

Mitra, R., Habib, K. N., Siemiatycki, M., Keil, R., Bowes, J., Taylor, R., . . . Siu, B. L. (2020). From Insight to Action on Transportation for Post-Secondary Students in the GTHA: 2019 Transportation Survey Findings. *StudentMoveTo*, 1-7. Retrieved October 28, 2020, from http://www.studentmoveto.ca/wp-content/uploads/2020/10/StudentMoveTO-2019-Report-Final-5-Updated-October-15-2020.pdf.

Webometrics Ranking of World Universities. (2020). Canada. Retrieved from https://www.webometrics.info/en/North_america/Canada.

List of Appendices

Screenshot of Code:

```
> # We uploaded the three csvs as ATT, HOUSE, and RES, and then merged them into
  a single df #
> merge1 <- merge(ATT, RES, by = "UniquePsKey_")
> merge2 <- merge(merge1, HOUSE, by = "UniquePsKey_")
> MTO <- merge2
>
> # Process for removing NA values from data #
> MTO_subset <- MTO[, c("psinstitution", "networkcommutedistance", "ps05CommuteF
actorInSchoolChoice", "psgender", "psage", "hhlivingsituation", "familyincomelev
el"),]
> MTO_clean <- MTO[complete.cases(MTO_subset),]
> MTO <- MTO_clean
> # Removing I don't Know and Prefer Not to Answer from select variables #
> MTO <- dplyr::filter(MTO, familyincomelevel != "I don't know", familyincomelev
el != "Prefer not to answer", psgender != "Prefer not to answer")
>
> # Assigning numerical categories (magnitude) to the DV #
> MTO$psinstitution_code <- revalue(MTO$psinstitution, c("University of Toronto"
 = 1, "McMaster University"=2, "York University" = 3, "Ryerson University" = 4,
 "Sheridan College" = 5, "OCAD University" = 6, "Centennial College" = 7, "Mohaw
k College" = 8, "Ontario Tech University" = 9, "Durham College" = 10))
>
> # Changing categorical/nominal data into factor #
> MTO$ps05CommuteFactorInSchoolChoice.f <- as.factor(MTO$ps05CommuteFactorInScho
olChoice)
> MTO$psgender.f <- as.factor(MTO$psgender)
> MTO$hhlivingsituation.f <- as.factor(MTO$hhlivingsituation)
> MTO$familyincomelevel.f <- as.factor(MTO$familyincomelevel)
>
> # Adjusting reference level of select variables #
> MTO$hhlivingsituation.f <- relevel(MTO$hhlivingsituation.f, 2)
> MTO$familyincomelevel.f <- relevel(MTO$familyincomelevel.f, 3)
```

Model Fit:

```
> model_mk6 <- lm(psinstitution_code ~ ps05CommuteFactorInSchoolChoice.f + netwo
rkcommutedistance + psgender.f + psage + hhlivingsituation.f + familyincomeleve
l.f, data = MTO)
> summary(model_mk6)
```

Regression Output:

```
Call:
lm(formula = psinstitution_code ~ ps05CommuteFactorInSchoolChoice.f +
    networkcommutedistance + psgender.f + psage + hhlivingsituation.f +
    familyincomelevel.f, data = MTO)

Residuals:
    Min      1Q  Median      3Q     Max
-5.3589 -1.9946 -0.4907  1.0049  8.1426

Coefficients:
                                                            Estimate
(Intercept)                                                2.939e+00
ps05CommuteFactorInSchoolChoice.fYes                       3.067e-01
networkcommutedistance                                     3.858e-06
psgender.fMale                                             3.095e-01
psgender.fNon binary / third gender                        6.121e-01
psage                                                      3.489e-02
hhlivingsituation.fLive alone                             -8.062e-01
hhlivingsituation.fLive with host family or at friend's house  1.182e+00
hhlivingsituation.fLive with partner                      -1.172e-01
hhlivingsituation.fLive with roommates                    -1.511e-01
familyincomelevel.f$ 100,000 - 124,999                    -7.158e-01
familyincomelevel.f$ 125,000 - 149,999                    -8.556e-01
familyincomelevel.f$ 150,000 - 199,999                    -1.027e+00
familyincomelevel.f$ 200,000 +                            -1.271e+00
familyincomelevel.f$ 30,000 - 39,999                      -7.430e-02
familyincomelevel.f$ 40,000 - 49,999                      -4.585e-01
familyincomelevel.f$ 50,000 - 59,999                      -7.295e-01
familyincomelevel.f$ 60,000 - 69,999                      -5.034e-01
familyincomelevel.f$ 70,000 - 79,999                      -5.268e-01
familyincomelevel.f$ 80,000 - 89,999                      -6.688e-01
familyincomelevel.f$ 90,000 - 99,999                      -9.180e-01
familyincomelevel.fLess than $ 14,999                      5.778e-01


                                                            Std. Error
(Intercept)                                                 1.957e-01
ps05CommuteFactorInSchoolChoice.fYes                        7.440e-02
networkcommutedistance                                      1.911e-06
psgender.fMale                                              7.391e-02
psgender.fNon binary / third gender                         3.018e-01
psage                                                       6.938e-03
hhlivingsituation.fLive alone                               1.474e-01
hhlivingsituation.fLive with host family or at friend's house  2.650e-01
hhlivingsituation.fLive with partner                        1.598e-01
hhlivingsituation.fLive with roommates                      9.555e-02
familyincomelevel.f$ 100,000 - 124,999                      1.473e-01
familyincomelevel.f$ 125,000 - 149,999                      1.781e-01
familyincomelevel.f$ 150,000 - 199,999                      1.776e-01
familyincomelevel.f$ 200,000 +                              1.752e-01
familyincomelevel.f$ 30,000 - 39,999                        1.586e-01
familyincomelevel.f$ 40,000 - 49,999                        1.571e-01
familyincomelevel.f$ 50,000 - 59,999                        1.536e-01
familyincomelevel.f$ 60,000 - 69,999                        1.598e-01
familyincomelevel.f$ 70,000 - 79,999                        1.716e-01
familyincomelevel.f$ 80,000 - 89,999                        1.740e-01
familyincomelevel.f$ 90,000 - 99,999                        1.764e-01
familyincomelevel.fLess than $ 14,999                       1.562e-01
```

```
                                                          t value
(Intercept)                                               15.017
ps05CommuteFactorInSchoolChoice.fYes                       4.122
networkcommutedistance                                     2.019
psgender.fMale                                             4.188
psgender.fNon binary / third gender                        2.028
psage                                                      5.029
hhlivingsituation.fLive alone                             -5.469
hhlivingsituation.fLive with host family or at friend's house   4.461
hhlivingsituation.fLive with partner                      -0.734
hhlivingsituation.fLive with roommates                    -1.581
familyincomelevel.f$ 100,000 - 124,999                    -4.861
familyincomelevel.f$ 125,000 - 149,999                    -4.804
familyincomelevel.f$ 150,000 - 199,999                    -5.784
familyincomelevel.f$ 200,000 +                            -7.252
familyincomelevel.f$ 30,000 - 39,999                      -0.468
familyincomelevel.f$ 40,000 - 49,999                      -2.918
familyincomelevel.f$ 50,000 - 59,999                      -4.749
familyincomelevel.f$ 60,000 - 69,999                      -3.151
familyincomelevel.f$ 70,000 - 79,999                      -3.070
familyincomelevel.f$ 80,000 - 89,999                      -3.844
familyincomelevel.f$ 90,000 - 99,999                      -5.205
familyincomelevel.fLess than $ 14,999                      3.700

                                                          Pr(>|t|)
(Intercept)                                                < 2e-16 ***
ps05CommuteFactorInSchoolChoice.fYes                      3.81e-05 ***
networkcommutedistance                                    0.043549 *
psgender.fMale                                            2.86e-05 ***
psgender.fNon binary / third gender                       0.042576 *
psage                                                     5.08e-07 ***
hhlivingsituation.fLive alone                             4.73e-08 ***
hhlivingsituation.fLive with host family or at friend's house 8.31e-06 ***
hhlivingsituation.fLive with partner                      0.463198
hhlivingsituation.fLive with roommates                    0.113832
familyincomelevel.f$ 100,000 - 124,999                    1.20e-06 ***
familyincomelevel.f$ 125,000 - 149,999                    1.60e-06 ***
familyincomelevel.f$ 150,000 - 199,999                    7.73e-09 ***
familyincomelevel.f$ 200,000 +                            4.68e-13 ***
familyincomelevel.f$ 30,000 - 39,999                      0.639525
familyincomelevel.f$ 40,000 - 49,999                      0.003535 **
familyincomelevel.f$ 50,000 - 59,999                      2.10e-06 ***
familyincomelevel.f$ 60,000 - 69,999                      0.001636 **
familyincomelevel.f$ 70,000 - 79,999                      0.002153 **
familyincomelevel.f$ 80,000 - 89,999                      0.000123 ***
familyincomelevel.f$ 90,000 - 99,999                      2.01e-07 ***
familyincomelevel.fLess than $ 14,999                     0.000218 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.514 on 5336 degrees of freedom
Multiple R-squared:  0.06083,   Adjusted R-squared:  0.05713
F-statistic: 16.46 on 21 and 5336 DF,  p-value: < 2.2e-16
```