UNDERSTANDING COLLEGE MAJOR CHOICE DECISIONS IN CANADA

by

Noémie Schmiedel

An essay submitted to the Department of Economics

in partial fulfillment of the requirements for

the degree of Master of Arts

Queen's University

Kingston, Ontario, Canada

September 2013

Acknowledgements

The Social Sciences and Humanities Research Council (SSHRC), the Education Policy Research Initiative (EPRI) and Queen's University are gratefully acknowledged for their funding assistance for this research, as are Statistic Canada and the analysts at the Queen's Research Data Centre (RDC) and COOL-RDC for access to the YITS-B dataset. I would also like to thank Steven Lehrer for all the helpful comments and advices, Stephen Childs for providing code that help generate a portion of the dataset used in this study and the participants at annual 2013 Canadian Economics Association (CEA) Conference for their thoughtful comments and suggestions.

Table of Contents

Section 1: Introduction **1**
Section 2: Literature **4**
Section 3: Data & Summary Statistics **8**
3.1 Data 8
3.2 Earnings and Unemployment rates 10
3.3 Limitations 12
3.4 Summary Statistics 13
Section 4: Modeling **17**
Section 5: Results & Discussion **23**
5.1 Multinomial Logit Specification 23
5.2 Mixed Logit specification 30
5.3 Possible improvements and extensions 31
Section 6: Conclusion **33**
References **36**
Appendix: Tables **39**

Section 1: Introduction

One of the most important choices an individual must make during his life is his educational choice. This dictates what type of jobs will be available to him and intrinsically influences the level of monetary comfort and happiness achievable in his professional life. Our project was motivated by our interest in the significant difference found in labour market outcomes across educational choices, more specifically college major choices, and our interest in studying the reasons for choosing different types of education and the career consequences of those choices.

Knowing the effect of the choice of major on wages and unemployment can also help determine the types of education that should be provided and the demand for post-secondary education by field of study. With policy reports calling for a major increase in college graduates in sciences, technology, engineering and math, the demand for higher education affects many aspects of government policy and is of special interest in education and labour market decisions. This paper can help determine where the demand is most concentrated, where funding should go and how best to redirect unemployed individuals.

Our approach is motivated by new and innovative research which incorporates the use of other labour market variables to more traditional research on determinants of college majors. Empirical research has traditionally found that the choice of major has a large impact on earnings. Differences in returns to majors can rival the college wage premium altogether (Altonji et al. (2012)) and remain significantly important even after controlling for pre-college test scores, math ability, hours worked in jobs and selection into the labour force (see Finnie & Frennette (2003)). With growing interest in the college major choice, research in the field has known great advancement. We now know more about the influential components of college major choice decisions, including: academic preparation, ability, preferences, family background and labour market conditions, to name only a few. This paper is highly influenced by Blom's new research (2012), which finds that unemployment rates and wages are important determinants of college major choices. She finds that high overall unemployment rates can increase labour supply elasticity and change some student behaviors with respect to their major choice. A more extensive and comprehensive review of the literature can be found in Section 2.

Thus, the goal of the present research is, firstly, to look at the question of choices for higher education and improve our understanding of how students make college major decisions, including how labour market decisions affect these choices. That is, studying the extent to which student decisions are influenced by regional wages and unemployment rates specific to each field of study. Secondly, we explore how high school curriculum, family background, individual level behavioral characteristics and self-reported skills can affect major decisions. Lastly, we investigate heterogeneity by gender, immigration status and region.

Our analysis uses the Youth in Transition Survey (YITS-B) cohort B, provided by Statistics Canada, which consists of a large and representative longitudinal micro data study on students between the ages of 18 and 20 at the start of the interviews. The YITS-B offers many advantages such as: information collected through five cycles, spanning from 1999 to 2007 which provides enough time variation to estimate the effects of labour market variables, a large sample size, collected from across the country to allow a more in depth analysis, extensive information on students' college major choices and background characteristics, both of which allow for sufficient control variables to estimate the determinants of college major choice decisions employed in this paper. This dataset was complemented with wages and unemployment rates from the CANSIM database (Statistic Canada). This addition provided both wages and unemployment rates by sex, at national and regional levels by field of occupation, which are used as proxies for the major fields of study. Combining the two created a unique dataset tailor made to further our understanding of college major decisions.

Making use of the YITS-B, this study employs two complementary longitudinally specific empirical strategies, developed to reveal the impact of college major choices on labour market variables. To examine the impact of major-specific labour market variables, as measured by wages and unemployment rates, we aggregate individual choices and use longitudinal and cross-sectional variation induced by demand shocks to estimate a set of supply equations for each field of study. From this we estimate the model using a multinomial logit specification and a mixed logit model, which relaxes the IIA assumption implicit with multinomial logit estimators. Such strategies will use the cross-sectional and longitudinal variations caused by exogenous shocks affecting the demand at the time of major choice. This requires extensive information on students' college major choices and background characteristics as well as enough time variation to estimate the effect of labour market variables, which the combined YITS-B and CANSIM databases provide. Both models are explained further in Section 4.1 and 4.2.

Our analysis yields some important findings for the education and labour economics literature on the determinants of college major choice. Firstly, we find that in the Canadian context, unemployment and wages seem to show statistically significant effects which concur with the American literature on the subject. The unemployment rate effect is particularly strong and consistent over all the fields of study evaluated. We also see important sorting within majors and significant involvement of many of the control variables, notably the self-rated skills and the high-school levels of math and language. Finally, we report statistically significant and important gender differences in both the distribution and trends of the fields of study. Women seem to put more weight on wages than men, while males seem to put a greater focus then females on unemployment rates when it comes to choosing their college major. These results are corroborated with the use of the mixed logit specification, which provides a significantly better representation of the major choice decision by relaxing the IIA property. See section 5 for the full discussion on the discovered results.

Section 2: Literature

Based on human capital theory, investment in human capital through education is an important determinant of economic development and growth. Coupled with the rising returns to schooling observed in the last decades, research on human capital and thus, education has decisively increased. However, the standard economic literature on education focuses mainly on educational attainment and the returns to education. Much of the interest on these subjects is grounded in empirical findings which suggest that returns to investment in education are substantial (See Psacharopoulos (1994) amongst others). Studies have shown that higher education is linked to higher wages and income as well as other social benefits.

Recently, more and more attention has been turned towards college majors and why individuals choose different types of education. Many factors, including expected earnings (e.g. Montmarquette et al. (2002) among many others), risk (e.g. Chen (2004) and Heckman et al. (2003)) and the roles of ability proxies, performance and preferences have been investigated. A number of researchers chose to incorporate expected earnings in their college major choice models by generally assuming that individuals compare the various expected outcomes at their disposal and choose the option that maximizes their expected utility (see Altonji (1993) and Arcidiacono (2004)). The Altonji et al. (2012) paper stresses the importance of knowledge early in high school to shape the feasibility and desirability of particular educational paths in combination with preferences and initial skills. Preferences and innate ability have been linked to the financial return of completing certain types of schooling even if they don't enter into the wage equation given that they influence the likelihood of making particular choices in programs of study. This implies certain associated payoffs. Altonji et al. also demonstrated empirically that preferences play an important role.

Researchers have also documented that the evidence of sorting into academic majors based on academic preparation. Arcidiacono et al. (2012) indicate that, in California, science majors in different campuses have on average stronger credentials than their non-science counterparts. A few studies have also shown that course selection in high school influences college major decisions and wages later in life. Not surprisingly, Altonji (1995) and Levine & Zimmerman (1995) find that taking additional mathematics and science courses increase the probability of choosing a technical college major. Rose & Betts (2004) look at the type and level of mathematics taught in high school and find that the returns are larger for advanced math courses, particularly algebra and geometry. While, Grogger & Eide (1995) show that the returns to math ability have been increasing over the years. Controlling for high-school grades has also been studied.

Other papers focus on the link between college major and occupation. Robst (2007) found that occupation and the extent to which human capital is major-specific play an important role in earning differences. Individuals employed in a field unrelated to their studies suffer a wage penalty that varies by field. As expected, specifically defined fields have bigger penalties than fields that develop general skills. For example, studying nursing and changing fields will produce a larger penalty than for someone with a business degree. Related work by Bonnard et al. (2012) on understanding expectations of ex ante earnings, which underlie major choices, show that parental involvement helps students make better forecasts. Forecasting errors also decline when the father's occupation is more closely connected to the child's discipline of study. Parental education has also been studied by Altonji & Dunn (1996), amongst others, revealing without too much surprise that parents with higher levels of schooling increased the chance for their children of having higher levels of education.

Researchers have uncovered large gender differences in the choice of major, as well as their returns. These continue to persist especially with less aggregated data. An important proportion of this disparity between major compositions comes from a disproportionate fraction of women who choose majors with lower earnings potential (see Turner & Bowen (1999) and Joy (2003)). For example, Turner and Bowen show that while the gender gap in science enrollment declined, this was driven by a surge of women in biology and there has been no change in the gender gap in mathematical sciences. Furthermore, most of the studies report that a portion of the wage gap comes from college major choice. Altonji et al. (2012) report that “much of the gender difference in the standard deviation is due to men selecting majors at the high end of the earnings distribution (e.g, engineering), rather than gender differences in the dispersion of the major coefficients.” Similarly, Zafar (2009) decomposes gender difference into differences in beliefs and preferences and finds that most of the gender gap comes from differences in preferences, not beliefs about ability. His results show that women’s choices depend about twice as much on non-pecuniary attributes compared to roughly equal proportions for men. Another recent contribution to this literature comes from Black et al. (2008). Evaluating the gender gap among college graduates while accounting for major choices, the authors find that well-educated women in the U.S. earned approximately 30% less than their counterpart, However, only 18% of that gap remains when controlling for highest degree, highest degree major field, and age. Results varied slightly across race and ethnicity.

The choice of college majors in post-secondary education has a substantial impact on earnings. For example, Altonji et al. (2012) find that the gap in log wage rates between electrical engineering and general education is of 56.1%. When compared to the 57.7% difference between college graduates and high-school graduates, the differences in returns across PSE fields of study can rival the college wage premium. Finnie & Frenette (2003), among many others, also showed that what you study affects your earnings. Generally, engineering consistently commands a high premium (around 0.40 relative to education) usually followed by business and science. This difference in returns to major remains even after controlling for precollege test scores, mathematics ability, hours worked in jobs and selection for the labour force. Walker & Zhu (2010) use quantile regression methods to estimate effects beyond the mean. They find negative returns at the bottom of the distribution or even at the median in Social Sciences (outside of Law, Economics and Management) as well as in Arts and Humanities.

Typically, college choice models predict that students should only respond to relative wages. However, a more recent paper has evolved to include other labour market variables and uses unemployment rate combined with wages to understand college major choices under more general circumstances (see Blom (2012)). Blom (2012) reports that labour supply elasticities are higher when overall levels of unemployment are high. She also states that when unemployment rates are high, students are more likely to choose better-paying, more difficult as well as more math-science and business concentrated majors. Possible explanations include: general economic conditions, tighter financial constraints and/or more awareness of returns during recessions. As recent studies[1] on formation of subjective expectations have found, students are not particularly well informed on the specifics of their choice. Generally, students are not aware of their major-specific wages and unemployment rates because this information is not widely distributed. However, the average level of unemployment is commonly and frequently reported by the media, especially in times of above average unemployment rates, such as during a recession. This type of information can lead some students to modify their major choice even without the information on their specific field of study. Furthermore, studies have found that local unemployment rates can have an important effect on the level of education students will wish to acquire. Growing enrollment in post-secondary education has been linked with rising unemployment rates (see Betts & McFarland (1995) or Bedard & Herman (2008)). This may partially be due to a reduction of the opportunity cost of attending college but could also indicate

[1]See Wiswall & Zafar (2012), Arcidiacono (2004), Hotz, & Kang (2011), Stinebrickner & Stinebrickner (2011) amongst others

that general economic conditions can affect the type of education students might choose. Blom (2012) chose to incorporate unemployment rates in the general models of college major choice to capture and evaluate the potential importance of this particular labour market variable as a determinant of college major. We follow in her footsteps by applying a similar methodology to the Canadian context.

In Canada, very little research on the subject has been done. Almost all of the evidence in literature referenced above uses data from the United States or Europe. Yet, there are many differences in the trends of college major choices depending on the region or country studied. For example, Akbari & Aydede (2012) document that student enrollment in the discipline of economics rose sharply in many countries, but not in Canada. They postulate that this may be a result of the continued perception of Canadian students that a lower economic reward is associated with an economics degree. With the lack of Canadian empirical evidence, this study first fills this important gap. Secondly, we explore how high school course selection may influence an individual's college major upon entering post-secondary education. Lastly, we conduct the analysis on the full sample and explore how the relationships vary across observed subgroups such as by gender.

Section 3: Data and Summary Statistics

3.1 Data

The data used in this study comes from the Youth in Transition Survey (YITS) Cohort B, which was conceived by Statistics Canada in partnership with Human Resources and Skills Development Canada (HRSDC). The YITS was developed as a longitudinal micro-data survey designed to study the patterns and influences on major transitions in early life, particularly with respect to education. Created especially to aid policy and program development to deal with both short term and long term problems, information from this study can also help assess the effectiveness of programs, the best time of implementation and which individuals should be targeted.

The YITS-B is a biennial study. It includes 22,378 respondents across Canada between the ages of 18 and 20 on December 31, 1999. The subjects were interviewed firstly in April 2000 to collect information on the year 1999 and retrospectively for earlier years. Then, it conducted follow-up interviews every two years to capture the activities that happened following the prior interviews. Hence, the last interview in 2008 captured the activities during 2006 and 2007, which provides us with 5 cycles of observations on a multitude of subjects. Hence, the YITS-B is clearly well suited to follow the young adults in their post-secondary education experience. The longitudinal aspect of the YITS allows for an evolution of the decisions and the possibility of changing such decisions along the way. It also provides sufficient variation over time to estimate the effects of the labour market variables, which is crucial for our analysis.

The YITS-B is representative of the national population of young adults making education choices as they tend to make post-secondary education choices around that age. Cohort B focuses on early post-secondary experiences with additional information such as gender, age, immigration status, financial information, high school experiences, family and parental background, skills and academic and social engagement. The YITS-B tracks a wide range of Post-Secondary Education (PSE) information within and over cycles from field of study to school status, switching or dropping out. This creates a digital map of their every decision at the post-secondary level. To the best of our knowledge, the YITS-B is the only Canadian survey that allows this type of analysis. The important number of observations in concordance with the range of variables offered by the YITS-B helps improve the meaningfulness and level of detail of our analysis of educational decisions at a comprehensive level, by field of study.

The data was supplemented with information from the Canadian socioeconomic database (CANSIM) from Statistics Canada which provided the two labour market variables: wages and unemployment rates by field of occupation. These variables were matched by time of high school graduation and used as proxies for the expected regional wages and regional unemployment rates for each field of study.[2] These major fields of studies were presented using the Classification Instructional Programs (CIP) or the Census Major Field of Study Classification (MSF), which we regrouped into twelve categories: ten sub-categories of PSE[3], a "Without PSE" category and a "dropout before graduating" category. For the analysis in this present paper, individuals that chose their post-secondary education field of study as "other" were excluded as they couldn't be classified in the twelve chosen subgroups. Since these students were few and created distortion due to their lack of commonalities, they were screened out as outliers. The final sample included 22,329 individuals, consisting of 5 cycles of cohort B subjects as defined with respect to their field of study.

3.2 Earnings & Unemployment rates

The two labour market variables studied are: wages (through expected earnings) and unemployment rates (through their expected value too). The earnings and unemployment rates available in the CANSIM database represent the hourly wage and unemployment rate an individual can expect for a certain field of occupation at the time of their choice. The occupational fields and fields of study were matched by hand to the best of our abilities to ensure that they would correspond to those available at the time decisions were made. That is, these measures would be considered to be

[2] We also looked at the influence of aggregate wages and unemployment rates by fields of study and at the labour market variables by sex and field of study. The results obtained for the various models proposed in Section 4 didn't present significantly different results from the ones presented in Section 5 using the regional wages and unemployment rates. The regions studied were: Atlantic Provinces, Quebec, Ontario, Prairies Provinces and British Columbia.

[3] 1- Business, Management & Public Administration; 2- Education; 3- Art & Humanities; 4- Social & behavioural Sciences; 5-Biology & Biomedical studies; 6-Other Sciences, Technologies, Engineering and Mathematics; 7-Health related studies; 8-Production and Transportation; 9- Agriculture, Natural resources, Conservation and Parks; 10-General Services (Food industry, Daycare, etc.);11- Dropout of PSE before graduation; 12-Without PSE

predetermined. Thereby, we paired each individual with his expected wage and unemployment rate based on the labour market values of these at the time of high school graduation[4], which is the time when the individual chooses which field of study, if any, he will pursue. This was a decision based on the fact that the student will make his/her decision based on the available data at the time of graduation. For the individuals who chose not to pursue PSE, the expected wage used was their wage at HS graduation time.

By using the wages and unemployment rates expected at time of graduation instead of the individual's own wage at the end of his studies we circumvent two important problems. Firstly, students have been known to misestimate their potential major-specific wage by as much as 28%[5]. Secondly, since wages can reflect ability, students can overestimate or underestimate their future abilities (to be developed during their studies). We avoid these problems using the high-school graduation time as the basis for their estimations of future labour market variables, which generally coincides with the time of major choice.

The choice to represent labour market variables by wages and unemployment rates reflect the assumption that students can't perfectly predict their future wages, nor their probability of unemployment. This choice also helps avoid the integral simultaneity problem found when estimating demand and supply equations, because wages and unemployment rates are predetermined at the time the decisions are made. They can also be perceived as exogenous to the students' choice of major as the students will not be entering the labour force for some years after their choice. It is also important to stress that we will compare wages and unemployment rates across the various major choices at the provincial level. Thus, there is substantial heterogeneity in these measures which will be shown in the summary statistics (Section 3.4).

[4] For individual who didn't finish high school, we used the last time they were in school as their "expectation date"
[5] See Wiswall & Zafar (2011) amongst others

3.3 Limitations

Despite the many advantages of using the YIT-B and CANSIM data for estimating the extent of students' decisions influenced by labour market conditions, there are a number of limitations that should be noted. Firstly, the YITS-B has inconsistencies between cycles due to differences in methods and scales in cycle 1. To use the information found in cycle 1 we had to match by hand the MSF fields of study to the CIP classification which can increase the possibility of false matches and false results. This matching problem also occurred when creating the wages and unemployment rates by field of study since the CANSIM dataset are classified using the 2006 National Occupational Classification for Statistics (NOC-S), while the YITS-B uses the CIP classification for the fields of study. Both were matched to the best of our abilities. However, it creates little discrepancies that are noted in the results[6]. Secondly, there is some loss of information with ineligible programs and inconsistencies in the records in terms of reporting their PSE programs and related fields of study. By the nature of the survey, programs and other entries were carried forward from one cycle to the next, which caused inconsistencies when respondents denied having chosen a particular program in the previous cycle of interviews. We treated this by excluding these ineligible programs from the calculation which effectively right-hand censored all the students who became ineligible from those inconsistencies. However, problems of understating switching or leaving and overstating persistence rates might occur from this treatment since this will exclude a relatively higher number of switchers and leavers.

The last limitation of YITS-B comes from the loss of information due to attrition of the dataset over the studied period. Table 1 shows the distribution of respondents over the five cycles of the YITS-B. An important problem of attrition is uncovered with only 44% of the original sample available at the end

[6] Due to wages and unemployment rates including occupations that do not require PSE. E.g. retail wages are captured in the Business category which brings down its average wage lower than expected.

of the fifth cycle. However, looking at the composition of the sample by cycle shows that the individuals leaving the study are not particularly dominating any category of interest. They seem to keep similar percentages over college major choices throughout the cycles and the other variables are relatively constant over the 8 years of the study. Table 1 shows that over the attrition of the sample, females increase by less than 0.4%, while immigrants decrease by less than 1.5%. Low-income and dropout respondents don't have a specific trend over the five cycles. This leads us to believe that attrition is mostly random in our case. The attrition seems to come from a homogeneous group of individuals rather than from a certain type of individual, which leads us to believe that attrition isn't a particularly important problem in our case. Rather than restrict the analysis to the respondents who were present until the last cycle and risk creating the associated sample bias, we preserved the individuals until they attrited themselves from the sample, which treated them as right-hand censored.

3.4 Summary Statistics

Table 2 presents summary statistics for the individuals selected in the sample of the YITS-B. The full sample is divided into eleven groups based on their first field of study after high-school. This also includes a category for the individuals who chose not to pursue post-secondary education. A majority of the sample (67.5%) have chosen post-secondary education while 32.5% have chosen not to pursue PSE. From the individuals who chose to pursue a PSE degree, the most popular fields of study are, in order: Business & management (14%), Arts & Humanities (13.1%) and STEMs other than Biology & Biomedical fields[7] (12.7%), while the least popular are: Education (1.8%), Primary sector (2.8%), Services (3%) and Biology & Biomedical (3.1%). This shows how certain fields of study can be more popular, while others have limited enrollment (e.g. biomedical).

[7] From now referred to as STEMS-B

Some important differences in the composition of those categories are to be noted. Firstly, female subjects are more highly concentrated in the education and health sectors with respectively 88.1% and 85.3% of these fields. However, women seem to have very low participation rates in production & Transport (8.9%) and STEMS-B (24.9%). These results could reflect a certain gender bias related to traditional roles. However, the knowledge of this divergence of choice can create opportunities for public policies who wish to increase the number of students in the STEM categories. As women seem to be underrepresented in these fields, appealing to the feminine population would be critical in gaining more growth in these sectors. This particular column also reveals an interesting fact: women are more present than men in post-secondary education. Only 40.5% of the females choose not to pursue PSE compared to the 59.5% of men. This implies that more and more women are getting higher education and that women are now more present then men at the post-secondary level. This is true despite the fact that the sample has slightly more men than women.

As for immigrants, they seem to target the more lucrative fields of study, with higher participation in STEM-B (14.2%), Business & Management (11.4%) and Biology & biomedical (11.2%). Individuals with partners (married or common law) at their time of major choice are more likely found in education (14%), Services (12.5%) or production (11%). This can show the more likely parings as women are highly dominating in education and men are highly dominating in production. Women who are in education are also more likely to be family oriented, which could explain this peak for married individuals.

In terms of income, a more important portion of the population seems to have a lower income[8] at the time of their college major choice (on average more than 40%). However, high and low income variables are complementary and mutually reinforcing. Individuals in Production and Transports have

[8] Low-income respondents refers to individual who had yearly income lower than the Low-Income Index (LII) gradually increases with age, starting at 7000$ ending at 25000$.

the highest percentage of high-income (9.7%) at time of choice and the lowest percentage of low-income (29%), while students in Health seem to be subjected to the inverse (1.7% & 45.5%). These trends can show where students need the most monetary help and where to allocate grants and scholarships to benefit the less fortunate.

As expected, individuals who chose not to pursue post-secondary education have on average the lowest wages and highest unemployment rate, respectively 8.3$/h. and 11.7%. Production, primary sector and services follow with above average unemployment rates and below average wages. The most remunerative fields with the lowest unemployment rates are unsurprisingly both of the Sciences driven categories: Biology & Biomedical and the Other Sciences, Technologies, Engineering and Mathematics (STEM-B) with respectively 17.7$/h & 5.8% and 16.9$/h & 6.5% unemployment. These results concur with previous literature.

One surprising outcome was the fact that Education is tied for highest wage at 17.7$/h. This is probably due to the inclusion of professors in the education category, which would bias the scales in favor of more lucrative wages than the average student in education can hope for. The same reasoning can explain the low wage for Business & Management, when it is known that they are one of the most profitable fields of study. The addition of retail and wholesale in this category's wage could have biased downward the hourly wage. This can be a problem as wages will not be reflective of the true potential of these two fields of study.

The table also shows the mean high school math grade average. Here sharp differences can be noted between Science related fields and the other PSE majors. With math averages above the 80% for both biology and STEM-B, we can safely assume that individuals who pursue studies in the STEMs fields of study have stronger mathematical skills than their counterparts. In contrast, individuals who chose not to pursue post-secondary education had the lowest math average with 72.7%. These results show

how strong logic skills and/or hard work in high school can translate into more interest in higher levels of education.

On a similar note, the compositions of switching and dropout activities have interesting elements. The highest level of switching occurred in Biology and Biomedical with roughly 31% of the students switching from this field to another. However, they have the lowest dropout rates. This implies that individuals in these fields change their field of study more frequently but still end up graduating. This could be to switch to less difficult majors for biomedical students. The STEMS-B fields have a very different story with above average dropout rates (12.7%) and below average switching (17.7%). This could be attributed to the fact that individuals who have affinities with sciences might be more inclined to switch to other sciences rather than radically change to social sciences or humanities. The switching distribution also reveals that Production and Services have the lowest rates of switching with respectively 7% and 9.4%. This could be attributed to the fact that these fields are broader, so interested individuals have more room to switch while remaining within them. On the other hand, the dropout rates are lowest for Biology & Biomedical (6.3%), Primary Sector (8.5%) and Education (8.6%). These three fields seem more restricted and specialized than the others. Higher switching and dropout rates can be a good indicator of peoples' abilities, since lower performance can be linked to higher dropout rates and switching to more manageable majors.

Overall, the summary statistics show that individuals choosing not to follow PSE have distinctively different characteristics. Moreover, there is important heterogeneity inside the PSE categories, which implies that a variety of factors can influence college major choices.

Section 4: Modeling

In this paper, our primary goal is to improve our understanding of college major decisions through the use of labour market variables, pre-college experience and other background variables such as influencing agents and controls. To do so, we used the cross-sectional and longitudinal aspects of our dataset to estimate two specifications: a multinomial logit and a mixed logit model. The standard approach used for college major choice is an individual discrete choice model, using either a conditional logit or multinomial logit model. However, the mixed logit specification has been getting growing attention in this area since it does not impose IIA condition, explained with further details shortly[9].

Following research by Blom (2012) and Altonji et al (2012) on labour market determinants of college major choices, we started with a simple expected utility equation where the *i*-th student is faced with a choice between the M major fields of studies offered to him. Suppose that the expected utility of choice *m* in region *r* is:

$$EU_{imr} = \eta_0 + \eta_1 wage_{mr} + \eta_2 unemp_{mr} + \alpha X_{ir} + \varepsilon_{imr} \tag{4.1}$$

where $wage_{mr}$ is the expected major-specific is hourly wage in region *r* and $unemp_{mr}$ is the associated level of unemployment at time of high school graduation. X_{ir} is subject *i*'s individual characteristics in region *r*, and ε_{imr} is the individual-specific shock. X_{ir} contains multiple dummy variables: male, immigrant, visible minority, low and high income, as well as multiple categorical regressors: parental education, parental occupation, number of siblings, income, number of dependent children, high school grade average and grade level for math & language, individual level behavioral characteristics and self-reported skills.

[9] In our case we clustered both specifications over region and year and frequency weighted by number of subjects in each field of study as to comply with Statistics Canada's requirements and to accommodate the mixed logit specification.

The student then makes their major choice between the $M\epsilon[1,12]$ available choices. If the student makes choice *m* in particular, then we can assume that his expected utility EU_{imr} is maximized among the M choices. The multinomial logit model is driven by the probability that choice *m* is made, which is that $\Pr\ (EU_{imr} > \ EU_{ikr})\ \forall\ k \neq m$.

Assuming a type I extreme-value distribution for the disturbances ε_{imr} [10], the probability of individual *i* choosing major *m* is:

$$\Pr\ (Y_i = m) = \frac{e^{\eta_0^{(m)}+\eta_1^{(m)}wage_{mr}+\eta_2^{(m)}unemp_{mr}+\alpha^{(m)}X_i}}{\sum_{j=1}^{12} e^{\eta_0^{(j)}+\eta_1^{(j)}wage_{jr}+\eta_2^{(j)}unemp_{jr}+\alpha^{(j)}X_i}} \tag{4.2}$$

Which can also be written as:

$$\Pr\ (Y_i = m|\, X_{ir}) = \frac{e^{\eta_0^{(m)}+\eta_1^{(m)}wage_{mr}+\eta_2^{(m)}unemp_{mr}}}{\sum_{j=1}^{12} e^{\eta_0^{(j)}+\eta_1^{(j)}wage_{jr}+\eta_2^{(j)}unemp_{jr}}} = \frac{e^{Z\eta^{(m)}}}{\sum_{j=1}^{12} e^{Z\eta^{(j)}}} \tag{4.3}$$

where *Z* is the vector of characteristics and $\eta^{(m)}$ is the vector of coefficients for choice *m*.

Hence, for this specification, we estimated the model for each of the m major fields of study proposed ($m\ \epsilon\ [1,12]$) using the category of "Without Post-secondary education" as the base outcome. This way we obtain the eleven multinomial logit equations which compare each of the categories 1, 2,…, 11 to the 12th category of "Without PSE".

The multinomial model generates a set of coefficients that capture the marginal effect of each regressor on the relative risk of each outcome relative to the base outcome (i.e, Without PSE). The regression results are displayed in terms of relative risk ratios (rrr) which have a relatively straight-forward and intuitive interpretation.

[10] Based on McFadden, D. (1973)

The relative risk ratio follows from the principle of relative risk, which simply put is the relative risk of an outcome, let say *m*, to the base outcome which is 12-Without PSE[11]. The relative probability or relative risk is:

$$\frac{\Pr\ (Y_i = m)}{\Pr\ (Y_i = 12)} = \frac{e^{Z\eta^{(m)}} / \sum_{j=1}^{12} e^{Z\eta^{(j)}}}{1/\sum_{j=1}^{12} e^{Z\eta^{(j)}}} = e^{Z\eta^{(m)}} \tag{4.4}$$

The ratio of relative risk (rrr) for a one-unit variation of one of the base characteristics, such as wage, is then:

$$\frac{e^{\eta_0^{(m)}+\eta_1^{(m)}(wage_{mr}+1)+\eta_2^{(m)}unemp_{mr}+\alpha^{(m)}X_{ir}}}{e^{\eta_0^{(m)}+\eta_1^{(m)}wage_{mr}+\eta_2^{(m)}unemp_{mr}+\alpha^{(m)}X_{ir}}} = e^{\eta_1^{(m)}} \tag{4.5}$$

Hence, the relative risk ratio for a one-unit change in a variable is in fact the exponential of the corresponding coefficients from the multinomial logit. This, in turn, can be evaluated more easily than the coefficients that the multinomial logit gives out. [12] The models that we evaluated contained the field of study as the dependant variable, the independent variables being: wage by occupation by region, unemployment rate by occupation by region, high school variables, general characteristics (including gender, age, income, and immigration status), individual level behavioral characteristics as well as self-rated skills in multiple disciplines.

The multinomial model imposes multiple assumptions on the data generation process. It generally assumes that the data is case specific and that the dependant variables can't perfectly predict the independent variable for any case. We also have to assume collinearity between the explanatory variables to be low. In the case of a choice model, the most important assumption is the independence

---

[11] By using category 12 as the base outcome, it effectively reduces $\eta^{(12)}$ to zero, as it will be used as the reference to which we will compare all other outcomes.
[12] Another possible way of threating the multinomial logit would be to present the marginal effects. However, with the sheer amount of observations and variables in our model, this would be time consuming and would not yield different results. They would simply be presented in a different units.

of irrelevant alternative assumption (IIA), which implies that the odds of preferring one choice over another don't depend on the presence/absence of other "irrelevant" alternative choices. For example, under the IIA, the relative probabilities of choosing a major in physics or economics should not change if a major in literature is added as an additional possibility. This is an important and restrictive assumption which can cause multiple problems if it is not respected. It is known that individuals often violate this assumption when making choices. The addition of choice might change things if the additional choice is a perfect or partial substitute for one of the evaluated choices. A problem might also arise due to indifference between choices or if the additional choice is considered better. In our case, imposing the IIA condition could be considered too restrictive and is likely violated. Thus, the multinomial logit model might be imposing an assumption too strong on the relative preferences and be considered undesirable.

This brings us to its alternative specification: the mixed logit model, also called the random parameters logit model. The mixed logit model is an extension of the multinomial model that currently acts as the modern way of modeling college major choices. It also examines discrete choices. However, it removes some limitations of the multinomial logit by relaxing the IIA property. By removing the imposition of the IIA, the mixed logit is better able to approximate adolescent decision making and is a more general model. The key benefit is that it also allows for random preference variations and correlations in unobserved factors over time, as well as explicitly accounting for individual correlations of unobserved utility in repetitive choices. Finally, the estimator and the distribution of coefficient are not restricted to a specific parametric distribution. Taken together, these properties allow the model to represent more general substitution patterns.

Based on the Revelt & Train (1998) and Train (2009) paper, the mixed logit estimator is underlined by a simple economic choice model where individuals aim to maximize expected utility similar to the model presented in 4.1. In this setting each individual is faced with 12 alternatives (the 12 possible fields of study) in each of the five cycles studied. Hence, we have the expected utility obtained by individual *i* from alternative *m* ($m \epsilon [1,12]$) in choice situation t ($t \epsilon [1,5]$):

$$EU_{imt} = \beta_{0i} + \beta_{1i}wage_{imt} + \beta_{2i}unemp_{imt} + \alpha_i X_{imt} + \varepsilon_{imt} \equiv \beta_i' z_{imt} + \varepsilon_{imt} \qquad (4.6)$$

where β_i is the individual-specific coefficients vector, z_{imt} is the vector of observed characteristics of individual *i* for alternative *m* in cycle t and ε_{imt} is the related error term. The error term is again assumed to be an independently and identically distributed extreme value.

This modification from the multinomial logit model adds one particularly important change to the basic model which is its random coefficients formulation. By allowing certain coefficients (β_i, vector of individual-specific coefficients) to be different for each individual, the mixed logit allows for random parameters. This allows the various coefficients to have different distributions. The unconditional probability of the observed order of the choices for individual *i* is given by the following formula:

$$P_i(\theta) = \int S_i(\beta_i) * f(\beta|\theta)\, d\beta \qquad (4.7)$$

where $f(\beta|\theta)$ is the density for a particular β and θ are the parameters of the distribution. $S_i(\beta_i)$ is the conditional probability of the choices if β_i is given. The proper formulation would be:

$$S_i(\beta_i) = \prod_{t=1}^{5} \frac{e^{\beta_i' z_{im(i,t)t}}}{\sum_{j=1}^{12} e^{\beta_i' z_{ijt}}} \qquad (4.8)$$

Where *m(i,t)* is the alternative that individual *i* chose in cycle *t*. Here, we can see where the multinomial logit formulation comes into play as it is used for the conditional logit specification $S_i(\beta_i)$. Hence, the mixed logit is sometimes viewed as "a weighted average of a product of logit formulas evaluated at different values of [β], with the weights given by the density $f(\beta|\theta)$."[13]

The goal of the mixed logit is to estimate the parameters of distribution of β which comes down to estimating the θ, the parameters that describe the distribution of the individual-specific coefficients. Thus, it estimates the mean and standard deviation of each individual's coefficients through a series of repeated draws of β_i, which are then used to get the $S_i(\beta_i)$. (see Train (2009) for more details). This implies that each individual can have different betas coming from this distribution. In general, the individual-coefficient vector and the corresponding expected utility can be presented as:

$$\beta_i = b + \gamma_i \qquad \& \qquad EU_{imt} = b' z_{imt} + \gamma' z_{imt} + \epsilon_{imt} \qquad (4.9\ a) \&\ b))$$

where *b* is the mean of the β_i and γ_i is the individual standard deviation, which represents the personal response difference with respect to the average response in the sample. The addition of the standard deviation portion allows for correlation over alternatives and cycles. This heterogeneity of the parameters explains why the IIA property does not need to hold in the mixed logit context.

The simulation based estimator employed by the mixed logit model is an approximation of the log likelihood of the model using simulation methods. In our model, we use a default of fifty Halton draws for the simulation. The simulated log likelihood is given by:

$$SLL(\theta) = \sum_{i=1}^{N} ln\left\{\frac{1}{50}\sum\nolimits_{h=1}^{50} S_i(\beta^h)\right\}$$

where *N* is the total number of individual, *h* represents the replications (in our case we used 50) and β^h is the *h*-th draw from the density function $f(\beta|\theta)$.

[13] See Hole, AR (2007) pp.389.

Overall, the mixed logit is a more general specification as it relaxes a number of assumptions implicit with the multinomial logit estimator and allows for individual and alternative-specific explanatory variables. It can also fit multiple models including the multinomial logit model as a special case. The results found with the mixed logit specification should therefore be more accurate than the ones found with the multinomial logit estimator. However, it is also more computationally complex. [14]

Section 5: Results & Discussion

5.1 Multinomial Logit Specification

Table 3 provides the multinomial logit relative risk ratio estimates of major coefficients by field of study. This specification uses category "12-Without post-secondary education" as the base outcome and evaluates the model over the five cycles, spanning from 1999 to 2007. The relative risk ratios of the multinomial logit are explained more thoroughly in Section 4. Table 3 shows the estimates of equation 4.5.[15] The full sample is presented in column one, while column two and three respectively present the subsamples for men and women. At the bottom of the table we also report the samples' size, pseudo-R-squared, number of iterations and the log likelihoods for each subsample.

Firstly, focusing on the wages by fields of study, we can see that for a one unit increase in hourly wage of the Other Sciences, Technologies, Engineering and Mathematics (STEMS-B) field the relative risk, relative to the Without post-secondary education category would be expected to increase by a factor of 1.183, holding all other variables constant. This means that the relative risk of studying STEMS-B fields is increased by roughly 18% compared to choosing not to pursue PSE, all else held constant. This

[14] In our case, with the size of the sample and the 5 cycle repetition, a full modelisation of the mixed logit specification would require computational power far superior than the one available to us in the RDCs at both the University of Ottawa and Queen's University.

[15] Standard interpretation of exponentiated coefficients uses odd ratios.

comes to confirm our hypotheses that increasing wages in a field should lead to more interest in this field, which implies increased interest in PSE.

Over the full sample, seven out the eleven categories lead to an increase in the relative risk (rrr>1)[16], four of these are statistically significant at a 0.1% level [17], the rest are statistically insignificant. On the other hand, four categories lead to a decrease in the relative risk compared to the Without PSE category (rrr<1)[18].) . Of these, only one is significant: the Dropout category. However, this category was mainly added to control for dropout individuals and observe the effects of certain variables on this choice. Yet, since most individuals don't enter post-secondary education with the idea of dropping out, the increase in the expected wage or unemployment rate should not particularly influence their choice.

In columns 2 and 3, wages seem to have a more pronounced effect on women than men. The relative risk ratios are consistently higher for women [19], with gender differences ranging from 2.6% to 15.3%. This seems to imply that an increase in hourly wage is important for women then for men when compared to the Without PSE category. This could be due to their historic role of secondary breadwinner. Women could typically stay at home without too much consequence. Hence, when it comes to work, they can be choosier. Raising their potential wage could increase their interest in post-secondary education. On the other hand, men are typically the primary breadwinner, which implies that they may focus more on finding a job rather than having a higher paying job. This could also partially explain the higher returns to schooling for women witnessed by countless papers (see Psacharopoulos (1994), amongst others).

Secondly, when looking at the unemployment rate results, we find that they are more important than the results for hourly wages. All the ratios of relative risk are under one (rrr<1) and they are all

---

[16] relative to the Without PSE category, all else held constant
[17] Education, Biology & Biomedical, STEMS-B and production
[18] Business & Management, Health, Services and Dropout of PSE
[19] Except for the Dropout category which we have explained, is not of particular importance for this variable.

highly significant (at a 0.1% level). This implies that if the unemployment rate were to increase by one unit, the relative risk of choosing to pursue a PSE in Education (for example), relative to not pursuing post-secondary education, would be expected to decrease by a factor of 0.693, the other variables in the model being held constant. All eleven categories produced similar estimates ranging from 0.536 to 0.77. These results indicate that students seem to attach great importance to the probability of finding work in their field of study. If their probability of finding work decreases in a particular field (which is equivalent to an increase in their expected unemployment rate), they will decrease their relative risk of studying in this field by at least 23% relative to not pursuing PSE, all other things being equal. This information corroborates our initial hypothesis that unemployment rate played an important role in determining college majors. A high unemployment rate in a field of study seems to decrease the interest of students in that field. This is an obvious conclusion since students want to be able to find work in their field following their studies. If unemployment rates are high in their field, it would be more difficult to do so. This would lead to a decreasing interest in the affected field.

Comparing column 2 and 3, we see that unemployment seems to matter more for men than women. Men have consistently lower relative risk ratios compared to women. This difference ranges from 3.1% to 11.3%. As an example, for men, an increase in unemployment rate by one unit is associated with a decrease in the relative risk for the field of Business (relative to not going to PSE) by a factor of 0.492. For women, this figure is 0.571. Both of these are significant at a 0.1% level. This implies that males react more strongly than females with regards to an increase in unemployment rates. Again, this could be explained by the role of first breadwinner males are typically imposed. Men might be less concerned about how much they earn and more about whether they earn compared to women. This would coincide with the findings of the previous wage discussion.

We also examined how a variety of other characteristics that could potentially affect the choice of field of study. Some clear and interesting patterns emerged. Firstly, looking at the High-school factors, we see that the relative risk ratios for high-school math and language levels are way above 1. Biology & Biomedical and STEMS-B have particularly high relative risk ratios in the math level case, while Education, Social Sciences and Arts & Humanities have higher relative risk ratio in the language level case. For example, a one-unit increase of the high-school math level will increase the relative ratio of the Business category by a factor of 1.82[20]. For the language level, this number drops to 1.34. Both are significant to the 5% level. Hence, for the field of Business & Management, math level is more important than language but both of them increase the odds of choosing to major in business rather than not pursuing post-secondary education. When comparing the estimates between men and women, we can't see any particular patterns emerging from these results, which is consistent with the academic literature but does differ with many popular anecdotes. On the other hand, when looking at the effect of the average grade, the relative risk ratio hovers around one, with little variation. The rrr interval for the full sample over the math grade average goes from 0.986 to 1.017, while for language it goes from 0.993 to 1.036. The overall grade average (GA) doesn't give more insight, with an interval from 1.01 to 1.064. This means that the grade average does not have an important effect on the choice of major, since an increase in the GAs by one unit will on average leave very little difference in relative risk between the fields and the without PSE category. This is somewhat surprising, since we would expect high school grade averages to be indicators of the student's abilities. However, this might be due to the fact that an increase in grade average by one unit is only an increase of one percent[21]. It would be more interesting to look at an increase of five or even ten percent to evaluate the importance of grade averages. Here again, There are no clear gender differences between the results and most of them are significant at a 5% level, with some small exceptions. In summary, the results suggest that what you study, and how you

[20] compared to the without PSE category, all else remaining constant
[21] It could also be attributed to the fact that self-rated skills would also measure ability.

perform in those courses may affect college major decisions but looking at more aggregated measures such as grade average masks these patterns. This may also suggest that students sort to courses based on their abilities beginning in high school.

After high-school factors, we look at individual level behavioral characteristics and self-reported skills. For the behavioral characteristics, the school activities participation [22] and number of friends pursuing PSE both have above one relative risk ratios (rrr>1), while most of the misbehaviour average indexes show relative risk ratios below one (rrr<1). The ranges of relative risk ratios for the full sample are respectively: [1.0; 1.375], [1.14; 1.242] and [0.626; 0.964].These results are significant for most of the fields of study but again show no particular difference between men and women in the subsamples. These results are expected. Additional school activity involvement and an increase in friends planning to further their education beyond high school are both signs of interest and engagement in school which should translate into a rise in relative risk of choosing a field of study[23]. To the opposite, an increase in the misbehaving index[24] is expected to lower this relative risk as it is an indicator of low interest and engagement in school. The results corroborate these hypotheses. As for the self-reported skills, we evaluated computer, writing and mathematics skills and found interesting results. The importance of skills depends on the field of study. For instance, the STEMS-B category has rrr>1 for computer and math skills, but rrr<1 for writing. On the other hand, Social Sciences has rrr>1 for computer and writing but rrr<1 for mathematics. This would indicate that individuals with particular sets of skills are more prone to choose related fields of study in order to utilize those skills to their full extent. These results could also imply that individuals tend to choose fields in which they have greater confidence in their abilities, as the variables were self-rated. When comparing these variables in columns 2 and 3, we find that self-

[22] This index measure the involvement in school related activities with an index ranging from 1 to 4, which rates how the individual liked to participate in school activities, for example clubs, sports, drama.
[23] relative to not pursuing a PSE, all else held constant
[24] Which takes into account alcohol consumption, skipping school and marijuana consumption. The higher the index (index range is from 1 to 5) the more misbehaving activities the individual has been doing.

reported skills with computers have more important effects for women and self-reported skills in writing seem to matter more for men. The effect of math skills is unclear for both genders but this may reflect the heterogeneity in math offerings across high schools. That is, a student in applied math may give a high self-report but the same student would provide a lower report in a more theoretical math course.

We also evaluated the effect a variety of general socioeconomic characteristics, including: gender, income and region. Here are some of the interesting results found in Table 3. For immigrants, the relative risk ratios are generally under one, ranging from 0.357 to 1.337 for STEMS-B[25], most of which are significant at a 5% level. This implies, for example, that immigrants relative to non-immigrants, the relative risk for Arts & Humanities[26] would be expected to decrease by a factor of 0.521. This would imply that immigrants are less likely than non-immigrants to choose these fields of study, relative to not pursuing PSE. Lower participation rates for immigrants could be attributed to many factors. One possibility would be the language barrier as more immigrants have to study in a language that is not their mother tongue. This would generate more schooling difficulties for these individuals which in turn would reduce their chances of pursuing PSE. The Atlantic, Quebec and Ontario regions all have similar effects hovering around the one relative risk ratio margin (rrr≈1). These are mostly statistically significant at conventional levels, respectively fluctuating between: [0.97; 1.0], [1.01; 1.05] and [088; 1.0]. This would imply that the region of residence at time of college major choice is not a particularly essential determinant of field of study. However, the high and low income dummy variables seem to have an important effect on the choice of major. The relative risk ratios are above one for low-incomes and below one for the high-incomes, varying between 1.055 - 1.912 and 0.024 - 0.569, respectively. Both are statistically significant at the conventional levels. This implies that for low-income

[25] The only other category that has an rrr > 1 is Business & Management. This could be due to the higher level of immigrants found in those fields, which was reported in section 3 using Table 2.
[26] compared to the Without PSE category, the other variables in the model being held constant

compared to not low-income, the relative risk ratio for STEMS-B[27] is expected to increase by a factor of 1.402. For the high-income relative to the not high-income, this factor drops to 0.347 and both are statistically significant at a 1% level. This would suggest that low-income individuals are more likely to pursue post-secondary education, while high-income individuals are less likely to do so. This goes against what we would typically believe but could still be explained by the fact that lower income individuals might have a lower income due to the fact that they are studying rather than making money working.

When comparing the estimates between columns 2 and 3 of Table 3, we don't see clear ordering between gender effects for most variables. However, we do see one for the low-income variable, where females seem to be more affected by the low-income compared to the not low-income. This would imply that lower income females are more preoccupied about getting PSE than their male counterparts. Another important difference between genders can be seen with the male status variable. The male status variable obtains a majority of relative risk ratios below the bar of one (rrr<1), except for the STEMS-B and the Production categories, which we have already established as the more male dominated fields of study. This implies that for males, compared to females, the expected relative ratio of the Business category (relative to Without PSE) should decrease by a factor of 0.492, all else held constant. This would imply that for most of the fields of study, women are more prone to follow post-secondary education than men. This ties-back perfectly with the summary stats found in Table 2.

Overall, the estimates presented in Table 3 suggest that a large number of factors seem to affect the decision making process of a student choosing his/her major. Most of the factors behave in a way we would expect ex-ante, but some of the estimated effects appear to be counter-intuitive. Of course, we observe some anomalies with the results, which can mostly be attributed to the data. Other

[27] relative to Without PSE, all else held constant.

measurement errors could also be the cause of some these incongruities. For example, expected wages are constructed by averaging over the occupational wages matching the field of study. However, many individuals do not end up working in their field of study which would create discrepancies between their expected earnings and their actual future earnings. The world of decision making is a complicated one as it is continually evolving with access to new information, as well as preferences changing over time. With the limited amount of data available to us, it's no wonder that it is difficult to construct a complete modeling that would take all of these aspects into consideration, which is why we see partial problems with the data. However, the main findings remain; wages and unemployment rates seem to significantly affect the choice of field of study, particularly unemployment rates.

5.2 Mixed Logit specification

Table 4 presents the results from the simplest mixed logit specification [28] using only wage and unemployment as random parameters affecting the choice of field of study. Due to the very large number of observations created with the wide form over the field of study, the analysis was cut to only evaluate the effect over cycle 2.[29] The simulation was performed using 50 Harlow random draws for each sampled field of study and we consider the normally distributed coefficients.

The estimated means and standard deviation of the coefficients presented in Table 3 provide information on the extent to which fields of study differ by showing the share of the sample that will place a positive or negative value on the variable. The distribution of the hourly wage finds an estimated mean of -0.158 and an estimated standard deviation of 0.202. This indicates that 78.3 percent of the

[28] More complete models were run but results were similar to the presented model and both take a lot more space to present and in certain circumstances have yet to be released from the RDC to ensure confidentiality is maintained.
[29] We chose to use cycle 2 as it has the closest estimates to the full model under the multinomial logit model. It also still has enough observations in the sample and limits the bias due to switching found in cycle 1.

distribution is above zero, while 21.7 percent is below[30]. This implies that higher hourly wage is a positive inducement for about three quarters of the students and a negative factor for the other quarter. On the other hand, the distribution of unemployment rates finds an estimated mean of 0.372 and an associated standard deviation of 0.473. This implies that roughly a quarter of the students prefer having a higher unemployment rate, while the other three quarters prefer having a lower unemployment rate.

The standard deviations of the random coefficients enters the model in a statistically significant manner, indicating that a mixed logit provides a significantly better representation of the major choice decision relative to the standard multinomial logit, which assumes that the coefficients are all the same for all majors. The factors that affect college major decisions have a heterogeneous relationship across the sample. However, even when we ignore this effect heterogeneity and estimate the multinomial logit model, a similar conclusion emerges: increasing hourly wages and reducing unemployment rates have a positive link on average to the decision of which college major to study. This also indicates why many information experiments that have been postulated by industry advocacy groups such as The Canadian Coalition for Tomorrow's Information and Communications Technology to boost attendance in ICT and STEM fields, by providing high school students with information on the labour market for graduates have had low returns. Our analysis shows that students seem to have a sense of this information, but the mixed logit results indicates there is a great deal of heterogeneity and future work needs to understand for which students this information could be valuable.

[30] These figures are given by $100 \times \Phi(-b_k/s_k)$, where $\Phi(-)$ is the cumulative standard normal distribution of $-$, b_k is the mean of the k-th coefficient (in our case hourly wage or unemployment rate) and s_k is its associated standard deviation.

5.3: Possible improvements and extensions

While these results corroborate the American findings on the subject, it might be interesting to change or extend those results to get even more precise results. Many improvements could be made in this regard. One simple improvement would be to extend the mixed logit model so it contains the full extent of the control variables, so we can see which variables vary across observations or have normally distributed coefficients. Another possible extension would be to change the classification of the fields. Due to the nature of the multinomial analysis, the classification was limited to twelve groupings. It would be interesting to redo this analysis with a larger and more detailed array of majors to see the precise effects of the major choice decision. This would also necessitate another dataset for the wage and unemployment rates as the CANSIM dataset is not equipped to provide sufficient detail for smaller classes.

Separating groupings such as Retail and Wholesale from Business & Management or professors from the Education category would provide more adequate estimates of the major choice decision. It would also be interesting to add in a factor that evaluates the expected level of post-secondary education. Other limitations of this study come from the choice of fields of study. Adding controls for switching or additional fields of study would give us a better perspective of the choices the students make when they have multiples field of study per cycle. In our case only one field of study could be chosen per cycle. This is necessary for statistical purposes; otherwise problems with the specifications could have arisen.

Section 6: Conclusion

As productivity is generally unknown by employers, the use of human capital, such as education, to evaluate the potential of a young worker is common. Hence, the importance of human capital and its link to higher wages. Higher education indicates potentially higher productivity, which leads to higher wages and more demand. As pecuniary and non-pecuniary compensations can vary greatly with respect to education and occupational choices, demand for and returns to types of human capital investment attract more and more interest from researchers and public policy analysts alike.

In this paper, we extend the Canadian education and labour economics literature on the determinants of college major choice by incorporating multiple labour market variables to the general model, specifically wages and unemployment rates. We also evaluate the complex model, which includes many controls, with two different specifications in order to have a clearer understanding of both the problem at hand and the public policy options available. We find that both wages and unemployment play an important role in determining college major choices. As unemployment in a field of study increases, students are more likely to choose other fields of study or not to pursue post-secondary education altogether. This effect is particularly strong in the men's case while the effect on wages is more important for females. This also shows the statistically significant and important gender differences in both the distribution and trends of the fields of study, as witnessed in the multinomial specification. Our study shows that men and women have very different patterns of selection when it comes to college majors. In addition, educational decisions are affected by numerous factors, some in an expected way, while others can have surprising effects. For example, we see that self-reported skills can have positive or negative effects depending on the chosen field of study. A specific instance of this would be that computer and writing skills affect positively Social sciences studies choices, while mathematical skills affect them negatively. An increase in mathematical skills by one unit (out of four) is expected to decrease the relative risk of choosing social sciences[31] by more than 25%. Therefore, we definitely see sorting within the fields of study. This paper also confirms a certain type of sorting at the high school level where we see that math and language level have an important and positive link with the college major decision as well as overall grades and skills. Therefore, influencing kids at a younger age could affect the sorting that we see at the high school level. A surprising result came from the effects of make skills where we would have expected a more significant overall effect and a clear difference between man and women. The insignificant results of gender difference in math skill can be attributed to various factors including heterogeneity in math offerings across high schools.

On the public policy front, these results can be very interesting as they can steer educational policies in the right direction. For example, while women have been increasingly choosing Biology & Biomedical fields, they remain vastly outnumbered by men in the more mathematical sciences. Hence, public policies focusing on boosting interest in STEM fields of study should appeal to a more feminine audience to improve their growth in a relatively unexploited section of the population. Through our study, we discovered that women are more likely to choose more mathematically intensive sciences (the STEM-B category) if they have a higher level of mathematics, more friends pursuing PSE or higher computer skills. A good plan of action would therefore be to provide free classes, tutors, perfecting sessions and skill boosting activities in math and computer sciences. Having extra access to computers and incorporating them in the class curriculums could also provide some help. Encouraging high-school students to choose more advanced levels of math by making math more fun at younger ages could also increase interest in math related fields. Our findings also show that low income students are more concentrated in STEM and that STEM students have higher grade averages, both overall and in math. Hence, allocation of scholarships and bursaries to lower income students and those with better academic achievements should also boost STEM post-secondary enrollment.

[31] Relative to the Without PSE category, all else held constant

Another potential plan would be to ensure better public access (read advertising here) to the average wages and unemployment rates of STEM graduates[32]. By providing insider knowledge on the determinants of college major choices, we can effectively influence student participation in those fields to meet job markets demands.

As we have stressed multiple times in this paper, we know that the factors affecting the college major choice decision have a heterogeneous relationship. This paper has listed numerous contrasting results between men and women. Women seem to be more affected by wage then men, while the latter are more preoccupied by unemployment rates than their counterparts. The use of the mixed logit specification has also confirmed presence of heterogeneity within the sample, where we see statistically significant and important returns on standard deviations of the coefficients. This implies that we can safely say that the effects of wage and unemployment rate can differ substantially over the sample. Students grant varying weights to these variables which proves the presence of heterogeneity. Understanding where this heterogeneity comes from is an important subject that needs to be accounted for and targeted in future research.

[32] As STEMs categories are some of the highest paid with the lowest unemployment rates in PSE, making this knowledge more accessible to students should increase their interest. Especially for women who react more strongly to increase in expected wages than men.

References:

Akbari, AH & Aydede, Y. (2012). Economic Benefits of Studying Economics in Canada: A Comparison of wages of Economics Majors with those in Other Disciplines Circa 2005. *CLSRN Working Paper, No. 109*.

Altonji, JG. (1993). The demand for and return to education when education outcomes are uncertain. *Journal of Labour Economics, Vol. 11 No.1*, pp. 48-83.

Altonji, JG. (1995). The effects of high school curriculum on education and labor market outcomes. *Journal of Human Resources, Vol. 30, pp.409–38.*

Altonji, JG & Dunn, TA. (1996). The Effects of Family Characteristics on the Return to Education. *The Review of Economics and Statistics*, *Vol. 78, No. 4 (Nov. 1996)*,pp. 692-704.

Altonji, JG, Blom E & Meghir C. (2012). Heterogeneity in Human Capital Investments: High School Curriculum, College Major, and Careers. *Annual Review of Economics, Vol. 4*, pp. 185-224.

Arcidiacono, P, Hotz, V & Kang, S. (2012). Affirmative Action and Sorting across Colleges: The effects of Prop 209 on Graduation Rates among University of California Campuses. Unpublished manuscript, Duke University.

Arcidiacono, P. (2004). Ability sorting and the returns to college major. *Journal of Economics , Vol. 121, No. 1-2, pp.343–75*.

Bedard, K., & Herman, D. (2008). Who goes to graduate/professional school? The importance of economic fluctuations, undergraduate field, and ability. *Economics of Education Review, Vol.27, Issue 2*, pp. 197-210.

Berger, M. (1988). Predicted future earnings and choice of college major. *Industrial and Labor Relations Review, Vol.41*, pp. 418-429.

Betts, J. &. (1995). Safe port in a storm: the impact of labor market conditions on community college enrollments. *Journal of Human Resources, Vol.30, No.4*, pp. 741-765.

Black, D., Haviland, A., Sanders, S., & Taylor, L. (2008). Gender Wage Disparities among the Highly Educated. *Journal of Human Resources, Vol.43, No.3*, pp. 630-659.

Blom E. (2012). Labor Market Determinants of College Major. *Working paper, MimeoYale University*.

Bonnard, C., Giret, J., & Mener, M. (2012). Educational intentions, cognitive skills and earnings expectations of French undergraduates. *Mimeo University of Lille*.

Chen, S. (2008). Estimating the variance of wages in the presence of selection and unobserved heterogeneity. *The Review of Economics and Statistics, Vol.90, No.2*, pp. 275-289.

Finnie R. & Frenette M. (2003). Earning differences by major field of study: evidence from three cohorts of recent Canadian graduates. *Economics of Education Review, Vol 22*, pp. 179-192.

Grogger, J. & Eide, E. (1995). Changes in college skills and the rise in the college wage premium. *Journal of Human Resources, Vol.30, No.2*, pp. 280-310.

Heckman, J, Carneiro, P. & Hansen, K. (2003). Estimating Distributions of Treatment Effects with an Application to the Returns to Schooling and Measurement of the Effects of Uncertainty on College Choice. *International Economic Review, Vol.44, No.2*, 361-422.

Hole, A. (2007). Fitting mixed logit models by using maximum simulated likelihood. *The Stata Journal, Vol.7, No. 3*, 388-401.

Joy, L. (2003). Salaries of recent male and female college graduates: Educational and labor market effects. *Industrial and Labor Relations Review*, 606-621.

Levine, P. & Zimmerman, D. (1995). The Benefit of Additional High-School Math and Science Classes for Young Men and Women. *Journal of Business & Economic Statistics, Vol.13, No.2*, pp. 137-149.

McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior. *Frontiers in Econometrics (Zaremmbka P. Ed.). Academic Press, New York.*

Monmarquette, C, Cannings, K. & Mahseredjian, S. (2002). How do young people choose college majors? *Economic of Education Review, Vol.21*, pp. 543-556.

Psacharopoulos, G. (1994). Returns to investment in education: A global update. *World development Vol.22, No. 9*, 1325-1343.

Revelt, D., & Train, K. (1998). Mixed logit with repeated choices: households' choices of appliance efficiency level. *Review of economics and statistics 80, no. 4*, 647-657.

Robst, J. (2007). Education and job match: the relatedness of college major and work. *Economics of Education Review, Vol.26, No. 4*, pp. 397-407.

Rose, H. & Betts, J. (2004). The Effect of High School Courses on Earnings. *Review of Economics and Statistics, Vol.86, No.2*, pp. 497-505.

Statistic Canada. *Table 282-0004 - Labour force survey estimates (LFS), by educational attainment, sex and age group, annual,* CANSIM (database). (accessed: 2013-06-24)

Statistic Canada. *Table 282-0070 - Labour force survey estimates (LFS), wages of employees by type of work, National Occupational Classification for Statistics (NOC-S), sex and age group, annual (current dollars) ,* CANSIM (database). (accessed: 2013-06-24)

Statistic Canada. (2008), Youth in Transition Survey (YITS) Cohort B, Canada.

Statistic Canada. *Table 282-0008 - Labour force survey estimates (LFS), by North American Industry Classification System (NAICS), sex and age group, annual,* CANSIM (database). (accessed: 2013-06-24)

Stinebrickner, T. & Stinebrickner, R. (2011). Math or science? Using longitudinal expectations data to examine the process of choosing a college major. *NBER working paper 16869*.

Train, K. (2009). *Discrete Choice Methods with Simulation, Chapter 6: Mixed Logit.* Cambridge University Press.

Turner, S. & Bowne, W. (1999). Choice of major: the changing (unchanging) gender gap. *Industrial and Labor Relations Review, Vol.52*, pp. 289–313.

Walker, I. & Zhu, Y. (2011). Differences by Degree: Evidence of the net financial rates of return to undergraduate study for England and Wales. *Economics of Education Review, Elsevier, Vol.30, No.6*, pp. 1177-1186.

Wiswall, M. & Zafar, B. (2011). *Determinants of college major choice: identification using information experience.* Staff Report 500, Federal Reserve Bank, New York, NY.

Zafar, B. (2009). *College major choice and the gender gap.* Staff Report 364, Federal Reserve Bank, New York, NY.

Appendix: Tables

TABLE 1: Distribution of respondents in the YITS-B surveys, 1999-2007

Cycle	observations	% of original sample	% of females	% of immigrants	% of low-income	% of dropouts
1	22 329	100.0	48.82	8.42	41.12	13.57
2	18 698	83.7	48.9	8.47	40.89	13.92
3	14 717	65.9	49.04	7.98	40.68	13.99
4	12 329	55.2	49.07	7.31	40.07	13.81
5	9 835	44.0	49.13	6.94	40.35	13.49

Note: Presenting distribution of respondent by cycle to evaluate level and severity of attrition. Analytically weighted with weights provided by the YITS-B. Percentage of females refers to the sex of the respondent. Percentage of immigrants refers to their status of landed immigrant. Percentage of low-income respondents refers to individual who had yearly income lower than the Low-Income Index (LII gradually increases with age, starting at 7000$ ending at 25000$). Percentage of dropouts refers to the individuals who started PSE but never graduated (sample reduced to only individual who took PSE).

Source: Statistics Canada YITS-B (1999-2007)

TABLE 2: College Possible Fields of Study Characteristics in the YITS-B surveys, 1999-2007

Fields	Share	% of females	% of immigrants	% of married~	% of low-income	% of high-income
1- Business & Mgmt.	14.0	59.2	11.4	8.5	39.3	3.3
2- Education	1.8	88.1	5.0	14.0	41.6	3.1
3- Art & Humanities	13.1	59.6	5.2	5.1	46.0	3.3
4- Social Sciences	7.1	72.5	7.6	6.7	38.7	4.0
5- Biology & Biomed.	3.1	65.5	11.2	3.3	35.8	3.8
6- Other STEM	12.7	24.9	14.2	5.0	44.0	4.0
7- Health	4.7	85.3	5.9	9.5	45.5	1.7
8- Production	5.2	8.9	6.1	11.0	29.0	9.7
9- Primary Sector	2.8	52.6	6.2	5.4	37.8	5.4
10- Services	3.0	57.9	4.0	12.5	40.5	2.9
12- Without PSE	32.5	40.5	7.6	10.4	40.9	7.3
Total	100.0	48.9	8.4	8.2	40.9	5.0

Fields	% of living in Ontario	Avg. Hourly wage[r]	Avg. Unemploy ment rate[r]	Avg. HS math grade	% of switchers	% of dropouts
1- Business & Mgmt.	42.1	12.8	7.0	76.5	16.5	13.5
2- Education	34.4	17.7	6.5	74.8	18.4	8.6
3- Art & Humanities	36.0	12.9	7.9	75.1	24.6	14.1
4- Social Sciences	42.8	13.9	6.9	74.1	27.1	13.4
5- Biology & Biomed.	44.8	17.7	5.8	83.2	31.1	6.3
6- Other STEM	36.8	16.9	6.5	80.9	17.7	12.7
7- Health	32.4	13.6	6.8	76.7	12.2	10.2
8- Production	32.1	11.6	10.0	74.5	7.0	14.2
9- Primary Sector	37.5	10.2	9.7	76.0	22.7	8.5
10- Services	48.7	10.7	8.5	74.2	9.4	12.5
12- Without PSE	32.2	8.3	11.7	72.7	-	-
Total	36.5	12.2	8.8	75.5	19.1	12.5

~ married or common law at time of choice of field of study

[r] Regional wages and Unemployment rates expected at time of high school graduation.

Note: Presenting Summary statistics to show heterogeneity of the composition of choices. Analytically weighted with weights provided by the YITS-B. 11 choices are shown (see detailed list in footnote -3), choice 11, composition is shown through % of dropouts. It shows how many individuals dropout before graduating PSE, which imply exclusion of category 12. Category 12 is also excluded for switchers, individuals who chose to change their general category of study. Share is the share of the total sample size found in each category. Low & high income are calculated at time of choice of field of study based on age of the individual (starting at 7000$ for 18-20 finishing at 20000$ for 26-28 years old).% living in Ontario based on residence at time of choice.

Source: Statistics Canada YITS-B (1999-2007)

<u>TABLE 3:</u> Ratio of Relative Risk of the multinomial logit regression of the of college major choice decision equation on selected fields of study with labour market variables, general, high school, skills and psychological controls. For 5 cycles respondents of the YITS-B (1999-2007)

Fields	RRR Full model	RRR Men	RRR Women
Business, Management & Public Admin			
Wage	0.958 (0.040)	0.927 (0.038)	0.990 (0.045)
Unemployment rate	0.536 (0.037)***	0.492 (0.039)***	0.571 (0.036)***
male status	0.492 (0.026)***	-	-
immigration status	1.184 (0.111)	1.026 (0.178)	1.254 (0.202)
visible minority	1.616 (0.146)***	1.616 (0.251)**	1.577 (0.202)***
Atlantic province status	0.976 (0.005)***	0.976 (0.009)**	0.976 (0.005)***
Quebec province status	1.065 (0.022)**	1.103 (0.029)***	1.037 (0.020)
Ontario province status	0.925 (0.013)***	0.854 (0.017)***	0.958 (0.020)*
Prairies province status	21.190 (18.220)***	32.109 (29.774)***	15.883 (12.377)***
parental occupation	9.796 (5.203)***	12.738 (7.477)***	8.286 (4.026)***
parental education	0.168 (0.118)*	0.155 (0.114)*	0.173 (0.116)**
siblings	0.089 (0.052)***	0.073 (0.047)***	0.107 (0.058)***
income (in 1000$)	1.036 (0.008)***	1.028 (0.007)***	1.066 (0.013)***
low income status	1.350 (0.195)*	1.423 (0.246)*	1.715 (0.227)***
high income status	0.215 (0.033)***	0.294 (0.040)***	0.145 (0.040)***
number of dependent children	0.569 (0.049)***	0.565 (0.062)***	0.521 (0.049)***
age group	1.905 (0.283)***	2.039 (0.294)***	1.799 (0.297)***
HS level of math	1.824 (0.177)***	1.929 (0.169)***	1.769 (0.213)***
HS level of language	1.344 (0.163)*	1.411 (0.146)**	1.303 (0.189)
HS math grade average	0.998 (0.002)	0.995 (0.004)	1.002 (0.003)
HS language grade average	1.014 (0.003)***	1.021 (0.006)***	1.006 (0.003)*

HS overall grade average	1.023 (0.007)***	1.018 (0.006)**	1.030 (0.007)***
school activities participation index	1.115 (0.034)***	1.174 (0.047)***	1.092 (0.045)*
number of friends going to PSE	1.186 (0.047)***	1.158 (0.048)***	1.225 (0.058)***
alcohol consumption (per month)	1.223 (0.041)***	1.260 (0.057)***	1.202 (0.048)***
misbehaviour index	0.851 (0.061)*	0.833 (0.068)*	0.861 (0.071)
self-reported skill with computers	1.801 (0.043)***	1.553 (0.072)***	1.955 (0.055)***
self-reported skill in writing	1.047 (0.028)	1.197 (0.035)***	0.911 (0.040)*
self-reported skill in mathematics	1.131 (0.026)***	1.146 (0.033)***	1.126 (0.033)***
constant	1.78E-07 (6.15E-07)***	7.62E-08 (2.39E-07)***	1.15E-07 (4.51E-07)***
Education			
wage	1.217 (0.035)***	1.132 (0.045)**	1.264 (0.045)***
unemployment	0.693 (0.034)***	0.636 (0.040)***	0.728 (0.031)***
male status	0.167 (0.019)***	-	-
immigration status	0.357 (0.135)**	0.509 (0.335)	0.304 (0.124)**
visible minority	0.463 (0.095)***	0.547 (0.288)	0.456 (0.117)**
Atlantic province status	0.975 (0.009)**	0.948 (0.025)*	0.975 (0.010)***
Quebec province status	1.060 (0.028)*	1.192 (0.056)***	1.028 (0.026)
Ontario province status	0.983 (0.024)	1.053 (0.063)	0.973 (0.029)
Prairies province status	10.974 (11.256)*	21.229 (25.407)*	8.239 (7.470)*
parental occupation	12.818 (7.797)***	28.404 (23.315)***	11.102 (5.883)***
parental education	0.112 (0.081)**	0.070 (0.077)*	0.112 (0.070)***
siblings	0.199 (0.130)*	0.205 (0.179)	0.224 (0.126)**
income (in 1000$)	1.049 (0.009)***	1.040 (0.008)***	1.081 (0.013)***
low income status	1.898 (0.299)***	2.428 (0.558)***	2.497 (0.420)***

high income status	0.024 (0.013)***	0.017 (0.028)*	0.023 (0.010)***
number of dependent children	0.504 (0.052)***	1.279 (0.258)	0.342 (0.050)***
age group	1.837 (0.281)***	1.899 (0.255)***	1.769 (0.315)**
HS level of math	1.389 (0.193)*	1.337 (0.274)	1.445 (0.217)*
HS level of language	1.988 (0.304)***	2.989 (0.969)**	1.929 (0.292)***
HS math grade average	0.994 (0.003)	0.982 (0.007)**	1.001 (0.003)
HS language grade average	1.012 (0.006)*	1.036 (0.010)***	1.000 (0.007)
HS overall grade average	1.036 (0.007)***	1.031 (0.011)**	1.041 (0.009)***
school activities participation index	1.375 (0.088)***	1.531 (0.177)***	1.365 (0.105)***
number of friends going to PSE	1.146 (0.060)**	1.566 (0.175)***	1.105 (0.064)
alcohol consumption (per month)	1.160 (0.056)**	1.033 (0.089)	1.204 (0.076)**
misbehaviour index	0.741 (0.048)***	1.051 (0.187)	0.679 (0.066)***
self-reported skill with computers	1.373 (0.048)***	1.090 (0.086)	1.431 (0.060)***
self-reported skill in writing	1.366 (0.101)***	1.854 (0.151)***	1.185 (0.104)
self-reported skill in mathematics	0.840 (0.045)**	0.944 (0.103)	0.825 (0.049)**
constant	1.17E-10 (4.14E-10)***	1.75E-14 (6.38E-14)***	7.23E-11 (2.76E-10)***
Art & Humanities			
wage	1.041 (0.038)	1.010 (0.034)	1.072 (0.045)
unemployment	0.630 (0.041)***	0.590 (0.045)***	0.664 (0.039)***
male status	0.682 (0.031)***	-	-
immigration status	0.731 (0.051)***	0.521 (0.085)***	0.903 (0.139)
visible minority	0.786 (0.090)*	0.688 (0.113)*	0.854 (0.131)
Atlantic province status	0.980 (0.005)***	0.986 (0.007)*	0.976 (0.007)**
Quebec province status	1.148 (0.022)***	1.157 (0.028)***	1.132 (0.021)***

Ontario province status	0.929 (0.020)**	0.933 (0.025)*	0.921 (0.030)*
Prairies province status	8.920 (7.969)*	10.658 (10.750)*	7.435 (5.850)*
parental occupation	8.620 (4.188)***	8.886 (5.411)***	8.133 (3.308)***
parental education	0.151 (0.116)*	0.130 (0.113)*	0.166 (0.114)**
siblings	0.108 (0.069)**	0.080 (0.061)**	0.138 (0.076)***
income (in 1000$)	1.013 (0.009)	1.004 (0.008)	1.040 (0.013)**
low income status	1.709 (0.231)***	1.963 (0.277)***	1.981 (0.245)***
high income status	0.227 (0.047)***	0.347 (0.081)***	0.140 (0.033)***
number of dependent children	0.320 (0.041)***	0.673 (0.120)*	0.232 (0.031)***
age group	1.657 (0.256)**	1.787 (0.266)***	1.555 (0.264)**
HS level of math	1.378 (0.159)**	1.379 (0.198)*	1.360 (0.153)**
HS level of language	2.012 (0.267)***	2.043 (0.167)***	2.012 (0.371)***
HS math grade average	0.994 (0.002)**	0.989 (0.003)**	0.998 (0.002)
HS language grade average	1.027 (0.003)***	1.033 (0.005)***	1.018 (0.005)***
HS overall grade average	1.044 (0.005)***	1.043 (0.004)***	1.049 (0.007)***
school activities participation index	1.201 (0.046)***	1.125 (0.044)**	1.268 (0.066)***
number of friends going to PSE	1.140 (0.050)**	1.147 (0.056)**	1.147 (0.054)**
alcohol consumption (per month)	1.043 (0.043)	0.981 (0.051)	1.097 (0.048)*
misbehaviour index	0.964 (0.045)	0.944 (0.074)	0.986 (0.042)
self-reported skill with computers	1.297 (0.046)***	1.189 (0.051)***	1.348 (0.053)***
self-reported skill in writing	1.518 (0.072)***	1.658 (0.076)***	1.380 (0.077)***
self-reported skill in mathematics	0.659 (0.015)***	0.633 (0.016)***	0.683 (0.021)***
constant	5.03E-09 (1.66E-08)***	1.09E-08 (3.19E-08)***	1.76E-09 (6.92E-09)***

Social Sciences			
wage	1.038 (0.026)	1.006 (0.027)	1.066 (0.031)*
unemployment	0.581 (0.028)***	0.549 (0.035)***	0.609 (0.025)***
male status	0.398 (0.020)***	-	-
immigration status	0.713 (0.091)**	0.798 (0.198)	0.657 (0.122)*
visible minority	1.415 (0.179)**	1.264 (0.240)	1.491 (0.228)**
Atlantic province status	0.977 (0.006)***	0.979 (0.011)	0.975 (0.007)***
Quebec province status	1.127 (0.026)***	1.163 (0.037)***	1.104 (0.021)***
Ontario province status	0.872 (0.025)***	0.804 (0.032)***	0.892 (0.030)**
Prairies province status	16.249 (14.112)**	15.586 (14.861)**	14.517 (11.352)**
parental occupation	6.476 (3.804)**	6.862 (4.504)**	5.978 (3.208)**
parental education	0.125 (0.099)**	0.076 (0.068)**	0.149 (0.108)**
siblings	0.084 (0.051)***	0.063 (0.046)***	0.103 (0.054)***
income (in 1000$)	1.028 (0.010)**	1.017 (0.009)	1.058 (0.014)***
low income status	1.620 (0.226)**	1.613 (0.302)***	2.114 (0.236)***
high income status	0.152 (0.027)***	0.122 (0.034)***	0.153 (0.040)***
number of dependent children	0.532 (0.082)***	0.547 (0.101)**	0.499 (0.083)***
age group	1.860 (0.301)***	1.979 (0.314)***	1.767 (0.313)**
HS level of math	1.594 (0.159)***	1.553 (0.182)***	1.611 (0.173)***
HS level of language	2.060 (0.220)***	2.589 (0.229)***	1.926 (0.268)***
HS math grade average	0.986 (0.002)***	0.972 (0.004)***	0.993 (0.002)**
HS language grade average	1.018 (0.005)***	1.023 (0.005)***	1.013 (0.005)*
HS overall grade average	1.035 (0.008)***	1.043 (0.006)***	1.035 (0.009)***
school activities participation index	1.186 (0.044)***	1.250 (0.065)***	1.188 (0.057)***

number of friends going to PSE	1.221 (0.050)***	1.513 (0.112)***	1.139 (0.049)**
alcohol consumption (per month)	1.194 (0.036)***	1.168 (0.057)**	1.220 (0.042)***
misbehaviour index	0.855 (0.043)**	0.930 (0.081)	0.825 (0.043)***
self-reported skill with computers	1.279 (0.041)***	1.205 (0.059)***	1.291 (0.041)***
self-reported skill in writing	1.417 (0.053)***	1.667 (0.074)***	1.249 (0.050)***
self-reported skill in mathematics	0.744 (0.024)***	0.750 (0.033)***	0.742 (0.028)***
constant	5.94E-09 (1.57E-08)***	1.45E-10 (3.31E-10)***	7.82E-09 (2.37E-08)***
Biology & Biomedical			
wage	1.243 (0.029)***	1.164 (0.022)***	1.317 (0.046)***
unemployment	0.693 (0.033)***	0.642 (0.037)***	0.738 (0.028)***
male status	0.421 (0.027)***	-	-
immigration status	0.867 (0.133)	0.468 (0.124)**	1.210 (0.189)
visible minority	1.802 (0.218)***	2.008 (0.385)***	1.651 (0.234)***
Atlantic province status	0.967 (0.008)***	0.973 (0.014)	0.962 (0.010)***
Quebec province status	1.149 (0.037)***	1.163 (0.042)***	1.134 (0.036)***
Ontario province status	0.948 (0.027)	0.885 (0.046)*	0.967 (0.031)
Prairies province status	11.703 (10.327)**	14.962 (14.341)**	9.628 (7.580)**
parental occupation	8.324 (5.399)**	11.280 (8.737)**	7.195 (4.080)**
parental education	0.075 (0.067)**	0.089 (0.086)*	0.065 (0.056)**
siblings	0.144 (0.086)**	0.124 (0.086)**	0.174 (0.088)**
income (in 1000$)	1.020 (0.010)*	1.012 (0.009)	1.049 (0.014)**
low income status	1.568 (0.246)**	1.616 (0.305)*	2.001 (0.313)***
high income status	0.205 (0.062)***	0.274 (0.096)***	0.132 (0.051)***
number of dependent children	0.338 (0.096)***	0.584 (0.217)	0.231 (0.066)***

age group	1.659 (0.344)*	1.390 (0.278)	1.765 (0.407)*
HS level of math	4.178 (0.473)***	3.987 (0.798)***	4.382 (0.622)***
HS level of language	1.455 (0.182)**	1.414 (0.193)*	1.533 (0.249)**
HS math grade average	1.017 (0.004)***	1.002 (0.006)	1.028 (0.005)***
HS language grade average	1.036 (0.004)***	1.036 (0.005)***	1.030 (0.006)***
HS overall grade average	1.064 (0.007)***	1.077 (0.013)***	1.062 (0.008)***
school activities participation index	1.300 (0.054)***	1.340 (0.079)***	1.320 (0.082***
number of friends going to PSE	1.242 (0.072)***	1.173 (0.073)*	1.290 (0.096)**
alcohol consumption (per month)	1.260 (0.073)***	1.252 (0.098)**	1.275 (0.082)***
misbehaviour index	0.760 (0.084)*	0.696 (0.092)**	0.798 (0.100)
self-reported skill with computers	1.277 (0.059)***	1.233 (0.083)**	1.319 (0.052)***
self-reported skill in writing	1.168 (0.044)***	1.267 (0.075)***	1.064 (0.065)
self-reported skill in mathematics	1.030 (0.052)	1.035 (0.071)	1.028 (0.056)
constant	5.64E-17 (1.13E-16)***	6.68E-16 (1.67E-15)***	1.97E-18 (4.81E-18)***
Sciences, Tech, Math & Engineer			
wage	1.183 (0.029)***	1.147 (0.024)***	1.275 (0.048)***
unemployment	0.674 (0.033)***	0.639 (0.040)***	0.736 (0.027)***
male status	1.745 (0.133)***	-	-
immigration status	1.337 (0.238)	1.314 (0.248)	1.076 (0.234)
visible minority	1.279 (0.204)	1.239 (0.213)	1.531 (0.322)*
Atlantic province status	0.984 (0.007)*	0.989 (0.008)	0.975 (0.009)**
Quebec province status	1.082 (0.025)**	1.091 (0.031)**	1.067 (0.024)**
Ontario province status	0.950 (0.018)**	0.958 (0.019)*	0.917 (0.029)**
Prairies province status	16.134 (12.234)***	21.572 (18.788)***	12.153 (7.228)***

parental occupation	9.794 (4.466)***	12.825 (6.944)***	6.727 (2.308)***
parental education	0.185 (0.131)*	0.208 (0.158)*	0.145 (0.098)**
siblings	0.199 (0.096)**	0.177 (0.099)**	0.251 (0.097)***
income (in 1000$)	1.018 (0.009)*	1.012 (0.008)	1.042 (0.012)***
low income status	1.402 (0.180)**	1.363 (0.199)*	1.791 (0.222)***
high income status	0.347 (0.061)***	0.412 (0.070)***	0.218 (0.068)***
number of dependent children	0.537 (0.054)***	0.607 (0.085)***	0.392 (0.080)***
age group	1.695 (0.284)**	1.766 (0.269)***	1.592 (0.324)*
HS level of math	2.234 (0.264)***	2.214 (0.272)***	2.253 (0.300)***
HS level of language	1.255 (0.143)*	1.303 (0.126)**	1.127 (0.208)
HS math grade average	1.015 (0.003)***	1.009 (0.004)*	1.025 (0.006)***
HS language grade average	1.012 (0.004)**	1.017 (0.005)**	0.997 (0.005)
HS overall grade average	1.035 (0.007)***	1.029 (0.009)**	1.056 (0.007)***
school activities participation index	1.000 (0.036)	0.961 (0.034)	1.117 (0.080)
number of friends going to PSE	1.212 (0.044)***	1.212 (0.044)***	1.213 (0.071)**
alcohol consumption (per month)	1.160 (0.039)***	1.171 (0.051)***	1.149 (0.063)*
misbehaviour index	0.817 (0.074)*	0.790 (0.082)*	0.882 (0.117)
self-reported skill with computers	2.235 (0.129)***	2.253 (0.131)***	2.144 (0.177)***
self-reported skill in writing	0.871 (0.031)***	0.939 (0.034)	0.767 (0.059)**
self-reported skill in mathematics	1.373 (0.051)***	1.354 (0.052)***	1.434 (0.075)***
constant	6.99E-11 (2.06E-10)***	2.75E-10 (8.41E-10)***	1.11E-11 (3.42E-11)***
Health			
wage	0.976 (0.051)	0.913 (0.071)	1.003 (0.050)
unemployment	0.536 (0.056)***	0.453 (0.089)***	0.566 (0.050)***

male status	0.184 (0.021)***	-	-
immigration status	0.779 (0.089)*	0.720 (0.194)	0.814 (0.120)
visible minority	1.196 (0.134)	2.049 (0.347)***	0.964 (0.131)
Atlantic province status	0.999 (0.009)	1.009 (0.020)	0.995 (0.009)
Quebec province status	1.061 (0.025)*	1.096 (0.034)**	1.042 (0.025)
Ontario province status	0.967 (0.024)	0.873 (0.034)***	0.983 (0.033)
Prairies province status	16.901 (17.265)**	42.927 (56.163)**	12.538 (11.164)**
parental occupation	8.363 (4.809)***	25.751 (18.638)***	6.325 (3.220)***
parental education	0.181 (0.149)*	0.262 (0.259)	0.162 (0.117)*
siblings	0.082 (0.063)**	0.116 (0.115)*	0.081 (0.055)***
income (in 1000$)	1.045 (0.008)***	1.033 (0.007)***	1.089 (0.014)***
low income status	1.912 (0.276)***	2.276 (0.454)***	3.002 (0.391)***
high income status	0.180 (0.044)***	0.310 (0.099)***	0.096 (0.029)***
number of dependent children	0.731 (0.092)*	0.330 (0.111)**	0.679 (0.079)**
age group	2.112 (0.312)***	2.103 (0.308)***	2.056 (0.342)***
HS level of math	1.854 (0.187)***	2.039 (0.302)***	1.829 (0.195)***
HS level of language	1.268 (0.163)	1.489 (0.231)*	1.262 (0.188)
HS math grade average	1.005 (0.003)	0.997 (0.007)	1.010 (0.004)*
HS language grade average	1.016 (0.003)***	1.031 (0.008)***	1.008 (0.004)*
HS overall grade average	1.033 (0.007)***	1.063 (0.011)***	1.027 (0.008)**
school activities participation index	1.219 (0.045)***	1.425 (0.067)***	1.209 (0.056)***
number of friends going to PSE	1.220 (0.052)***	1.271 (0.074)***	1.222 (0.059)***
alcohol consumption (per month)	1.200 (0.043)***	1.261 (0.079)***	1.193 (0.054)***
misbehaviour index	0.731 (0.044)***	0.661 (0.066)***	0.760 (0.054)***

self-reported skill with computers	1.151 (0.047)**	1.080 (0.079)	1.117 (0.051)*
self-reported skill in writing	1.134 (0.032)***	1.343 (0.092)***	1.022 (0.040)
self-reported skill in mathematics	1.003 (0.036)	1.082 (0.087)	0.990 (0.045)
constant	5.65E-07 (1.92E-06)***	4.21E-10 (1.90E-09)***	6.52E-07 (2.35E-06)***
Production			
wage	1.111 (0.024)***	1.101 (0.023)***	1.169 (0.035)***
unemployment	0.770 (0.051)***	0.751 (0.057)***	0.816 (0.039)***
male status	5.412 (0.829)***	-	-
immigration status	0.830 (0.160)	0.756 (0.155)	1.252 (0.485)
visible minority	0.683 (0.082)**	0.625 (0.112)**	0.886 (0.446)
Atlantic province status	0.989 (0.006)	0.985 (0.006)***	1.040 (0.019)*
Quebec province status	1.047 (0.022)*	1.038 (0.023)	1.177 (0.041)***
Ontario province status	0.926 (0.019)***	0.931 (0.023)**	0.902 (0.063)
Prairies province status	6.071 (5.064)*	7.172 (6.284)*	3.541 (2.970)
parental occupation	3.225 (1.741)*	2.919 (1.692)	16.061 (10.404)***
parental education	0.346 (0.237)	0.344 (0.246)	0.237 (0.177)
siblings	0.241 (0.164)*	0.212 (0.152)*	0.364 (0.258)
income (in 1000$)	1.024 (0.006)***	1.018 (0.005)**	1.059 (0.017)***
low income status	1.186 (0.136)	1.134 (0.158)	1.789 (0.324)**
high income status	0.569 (0.108)**	0.658 (0.112)*	0.269 (0.185)
number of dependent children	0.761 (0.055)***	0.793 (0.077)*	0.533 (0.144)*
age group	1.498 (0.170)***	1.509 (0.160)***	1.525 (0.266)*
HS level of math	1.433 (0.123)***	1.371 (0.119)***	2.656 (0.475)***
HS level of language	1.208 (0.124)	1.206 (0.124)	1.502 (0.278)*

HS math grade average	0.990 (0.003)**	0.987 (0.003)***	0.994 (0.006)
HS language grade average	0.993 (0.003)*	0.993 (0.003)*	0.996 (0.010)
HS overall grade average	1.014 (0.005)**	1.015 (0.004)***	1.005 (0.013)
school activities participation index	1.025 (0.050)	1.002 (0.051)	1.139 (0.118)
number of friends going to PSE	1.176 (0.038)***	1.163 (0.041)***	1.354 (0.160)*
alcohol consumption (per month)	1.204 (0.050)***	1.182 (0.049)***	1.562 (0.178)***
misbehaviour index	0.819 (0.047)**	0.835 (0.049)**	0.483 (0.092)***
self-reported skill with computers	1.018 (0.028)	0.970 (0.028)	1.608 (0.121)***
self-reported skill in writing	0.917 (0.032)***	0.936 (0.034)	0.974 (0.074)
self-reported skill in mathematics	1.300 (0.034)***	1.303 (0.032)***	1.311 (0.128)**
constant	1.84E-05 (3.71E-05)***	0.0004 (0.0008)***	4.69E-12 (1.55E-11)***
Agriculture & primary sector			
wage	1.022 (0.041)	1.002 (0.032)	1.048 (0.056)
unemployment	0.707 (0.047)***	0.695 (0.054)***	0.726 (0.043)***
male status	0.663 (0.050)***	-	-
immigration status	0.503 (0.130)**	0.167 (0.081)***	0.909 (0.272)
visible minority	0.612 (0.116)*	0.770 (0.123)	0.463 (0.147)*
Atlantic province status	0.996 (0.006)	1.004 (0.011)	0.989 (0.010)
Quebec province status	1.134 (0.017)***	1.127 (0.027)***	1.139 (0.017)***
Ontario province status	0.941 (0.026)*	0.931 (0.041)	0.947 (0.033)
Prairies province status	3.727 (3.262)	4.680 (4.368)	2.823 (2.226)
parental occupation	4.715 (2.257)**	3.338 (1.492)**	6.561 (3.416)***
parental education	0.180 (0.134)*	0.185 (0.144)*	0.162 (0.116)*
siblings	0.150 (0.099)**	0.154 (0.108)**	0.152 (0.096)**

income (in 1000$)	1.017 (0.008)*	1.004 (0.007)	1.048 (0.011)***
low income status	1.474 (0.183)**	1.305 (0.195)	2.035 (0.252)***
high income status	0.351 (0.075)***	0.431 (0.130)**	0.280 (0.082)***
number of dependent children	0.382 (0.055)***	0.689 (0.121)*	0.243 (0.050)***
age group	1.780 (0.275)***	1.873 (0.235)***	1.693 (0.337)**
HS level of math	1.932 (0.150)***	1.861 (0.234)***	1.988 (0.246)***
HS level of language	1.644 (0.167)***	1.375 (0.162)**	2.094 (0.248)***
HS math grade average	0.998 (0.003)	0.991 (0.004)*	1.004 (0.004)
HS language grade average	1.000 (0.003)	1.002 (0.006)	0.996 (0.004)
HS overall grade average	1.035 (0.007)***	1.045 (0.008)***	1.030 (0.009)***
school activities participation index	1.722 (0.110)***	1.802 (0.140)***	1.666 (0.118)***
number of friends going to PSE	1.170 (0.052)***	1.203 (0.071)**	1.148 (0.064)*
alcohol consumption (per month)	1.445 (0.067)***	1.499 (0.088)***	1.406 (0.068)***
misbehaviour index	0.626 (0.047)***	0.636 (0.055)***	0.610 (0.057)***
self-reported skill with computers	1.218 (0.050)***	1.180 (0.047)***	1.262 (0.067)***
self-reported skill in writing	1.016 (0.042)	1.055 (0.066)	0.970 (0.046)
self-reported skill in mathematics	1.047 (0.031)	0.990 (0.042)	1.102 (0.045)*
constant	2.33E-09 (6.30E-09)***	2.16E-08 (5.31E-08)***	5.34E-11 (1.89E-10)***
Services			
wage	0.956 (0.047)	0.948 (0.046)	0.974 (0.050)
unemployment	0.603 (0.048)***	0.581 (0.051)***	0.632 (0.048)***
male status	0.849 (0.123)	-	-
immigration status	0.675 (0.128)*	0.678 (0.147)	0.641 (0.202)
visible minority	0.914 (0.167)	0.956 (0.219)	0.855 (0.222)

Atlantic province status	0.999 (0.011)	0.986 (0.018)	1.008 (0.015)
Quebec province status	1.012 (0.020)	1.064 (0.030)*	0.962 (0.022)
Ontario province status	0.940 (0.021)**	0.992 (0.031)	0.895 (0.031)**
Prairies province status	10.251 (8.637)**	11.243 (11.324)*	9.444 (6.740)**
parental occupation	5.054 (2.706)**	5.839 (3.847)**	4.547 (2.276)**
parental education	0.343 (0.236)	0.387 (0.322)	0.317 (0.191)
siblings	0.093 (0.060)***	0.070 (0.059)**	0.131 (0.070)***
income (in 1000$)	1.005 (0.009)	1.001 (0.010)	1.024 (0.013)
low income status	1.055 (0.167)	0.852 (0.163)	1.445 (0.244)*
high income status	0.569 (0.192)	0.527 (0.190)	0.657 (0.295)
number of dependent children	0.626 (0.082)***	0.846 (0.142)	0.495 (0.085)***
age group	1.839 (0.291)***	1.889 (0.293)***	1.742 (0.295)**
HS level of math	1.207 (0.143)	1.387 (0.227)*	1.063 (0.112)
HS level of language	1.384 (0.224)*	1.297 (0.170)*	1.457 (0.314)
HS math grade average	0.999 (0.005)	0.993 (0.006)	1.003 (0.005)
HS language grade average	1.003 (0.006)	1.016 (0.007)*	0.990 (0.007)
HS overall grade average	1.013 (0.007)	1.008 (0.009)	1.021 (0.009)*
school activities participation index	1.366 (0.067)***	1.642 (0.127)***	1.212 (0.069)**
number of friends going to PSE	1.158 (0.052)**	1.257 (0.083)**	1.101 (0.071)
alcohol consumption (per month)	1.216 (0.054)***	1.231 (0.069)***	1.216 (0.072)**
misbehaviour index	0.795 (0.057)**	0.728 (0.066)***	0.872 (0.078)
self-reported skill with computers	0.987 (0.043)	0.960 (0.042)	0.979 (0.067)
self-reported skill in writing	1.179 (0.058)**	1.339 (0.067)***	1.028 (0.075)
self-reported skill in mathematics	0.860 (0.032)***	0.848 (0.042)**	0.879 (0.039)**

constant	0.001 (0.003)	0.0001 (0.0005)*	0.003 (0.011)
Dropout of PSE			
wage	0.833 (0.027)***	0.838 (0.023)***	0.829 (0.033)***
unemployment	0.565 (0.028)***	0.554 (0.032)***	0.573 (0.027)***
male status	1.043 (0.027)	-	-
immigration status	0.913 (0.063)	0.948 (0.104)	0.826 (0.153)
visible minority	0.909 (0.077)	0.845 (0.076)	1.047 (0.123)
Atlantic province status	0.991 (0.003)**	0.992 (0.004)	0.989 (0.007)
Quebec province status	1.047 (0.013)***	1.071 (0.019)***	1.013 (0.012)
Ontario province status	0.999 (0.009)	1.007 (0.021)	0.989 (0.016)
Prairies province status	10.415 (8.967)**	13.787 (12.361)**	8.340 (6.953)***
parental occupation	7.774 (4.087)***	9.146 (5.037)***	6.774 (3.483)***
parental education	0.134 (0.095)**	0.150 (0.110)**	0.120 (0.083)**
siblings	0.059 (0.036)***	0.060 (0.039)***	0.058 (0.035)***
income (in 1000$)	1.036 (0.008)***	1.030 (0.007)***	1.061 (0.013)***
low income status	1.181 (0.125)	1.054 (0.131)	1.548 (0.179)***
high income status	0.308 (0.057)***	0.343 (0.066)***	0.314 (0.076)***
number of dependent children	1.091 (0.064)	0.871 (0.063)	1.172 (0.075)*
age group	2.088 (0.215)***	2.111 (0.195)***	2.062 (0.256)***
HS level of math	1.347 (0.074)***	1.351 (0.078)***	1.332 (0.085)***
HS level of language	1.510 (0.171)***	1.453 (0.150)***	1.592 (0.212)***
HS math grade average	0.996 (0.001)**	0.993 (0.002)***	1.000 (0.002)
HS language grade average	1.011 (0.002)***	1.015 (0.003)***	1.006 (0.004)
HS overall grade average	1.010 (0.005)*	1.010 (0.006)	1.012 (0.004)**

school activities participation index	1.048 (0.030)	1.021 (0.026)	1.107 (0.049)*
number of friends going to PSE	1.143 (0.022)***	1.127 (0.034)***	1.167 (0.032)***
alcohol consumption (per month)	1.021 (0.023)	1.034 (0.029)	1.021 (0.035)
misbehaviour index	1.046 (0.037)	0.997 (0.051)	1.102 (0.049)
self-reported skill with computers	1.300 (0.033)***	1.290 (0.033)***	1.259 (0.052)***
self-reported skill in writing	1.074 (0.024)**	1.123 (0.029)***	1.016 (0.039)
self-reported skill in mathematics	0.961 (0.018)*	1.019 (0.017)	0.899 (0.032)**
constant	0.001 (0.003)**	0.002 (0.005)**	0.001 (0.002)**
Without PSE	**Base outcome**		
Log likelihood	-474603.5	-237553.3	-231761.0
Number of iterations	6	6	6
Pseudo R2	0.278	0.278	0.259
Number of Obs.	420154	230285	189869

Note: ***, **, * respectively denote statistically different from zero at the 0.1%, 1% and 5% confidence levels; robust standard errors in parentheses.
Presenting relative risk ratios (rrr) of the multinomial logit specification (see equation 4.5). Dependant variable is field of study chosen between 12 categories. Individual who didn't fit those categories were excluded. Using "Without PSE" as the base outcome. Observations are frequency weighted by the number of individual contributing to each field of study's category and clustered over cycle and region.
Source: Statistics Canada YITS-B (1999-2007) for most variables. Using regional hourly wage and unemployment rate from Statistics Canada CANSIM at time of high school graduation on selected fields of post-secondary education. Possible regions are: Atlantic Provinces, Quebec, Ontario, Prairies Provinces and British Columbia. Evaluated over 5 cycles.

<u>TABLE 4:</u> Mixed logit model of the of college major choice decision equation on selected fields of study with labour market variables, for cycle 2 respondents of the YITS-B (2000-2001)

Variables	**Parameter**	**Value**	**Robust Standard Error**
Hourly wage			
	Mean of coefficient	-0.158	0.021***
	Standard deviation of coefficient	0.202	0.017***
Unemployment rate			
	Mean of coefficient	0.372	0.040***
	Standard deviation of coefficient	0.473	0.035***
Log likelihood		-232131.05	
Number of iterations		6	
Number of Obs.		1594334	

<u>Note:</u> ***, **, * respectively denote statistical significance at the 0.1%, 1% and 5% confidence levels; Presenting mean and standard deviations of coefficients (betas) of the mixed logit specification (see equations 4.9 a) & b)). Dependent variable is a dummy variable that indicates if the field of study was chosen during cycle 2 from Statistics Canada YITS-B (2002). Observations are shaped in wide format over field of study. Independent variables are hourly wages and unemployment rates high school graduation on selected fields of post-secondary education from Statistics Canada CANSIM database (see section 3 for list). Observations are frequency weighted by the number of individual contributing to each field of study's category and clustered over cycle and region.

<u>Source:</u> Statistics Canada YITS-B (1999-2007) & CANSIM database (1991-2007)