MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„The Role of Identity in the Occupation Choice"

verfasst von / submitted by

Violeta Andzelina Kuzikaite, BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2020 / Vienna 2020

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	UA 066 913
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	MasterstudiumVolkswirtschaftslehre
Betreut von / Supervisor:	Univ.-Prof. Dipl.-Ing. Dr. Christine Zulehner, Privatdoz

Abstract

The purpose of this thesis is to analyse the underlying causes of gender gap through education selection choices and later career choices. The traditional economic analysis of the education choice lacks in explanation why highly skilled individuals choose low paid careers. This thesis incorporates the role of identity in the economic analysis of education choice. The study is based on the Swiss survey of Transitions from Education to Employment (TREE). Applying the factor analysis on a range of survey questions about the attitudes, the measures of identity are constructed. These measures are incorporated in the logit and multinomial logit analysis to study the education and career choice. The results indicate that gender differences persist in self-perception. The focus on self-identity may provide some explanation to policy makers on why financial incentives may be insufficient to attract individuals to some occupations with excess demand.

Abstract in German / Abstract auf Deutsch

Ziel dieser Arbeit ist, die Hintergründe des Gender Gaps, der aus der Auswahl der Bildungs- und späteren Karrierewegsentscheidungen resultiert, zu beleuchten. In der traditionellen ökonomischen Analyse der Bildungswegentscheidung werden die Beweggründe, warum begabte Individuen Karrieren in niedrig bezahlten Sektoren wählen, nicht hinreichend berücksichtigt. Dieser Text zieht auch die Rolle der Identität bei der Analyse der Karriereentscheidung in Betracht. Die Studie basiert auf der Swiss survey of Transitions from Education to Employment (TREE). Durch Anwendung der Faktorenanalyse auf eine Reihe von Fragestellungen der Studie hinsichtlich der Einstellungen werden Maßstäbe für die Identität konstruiert. Diese Maßstäbe werden in die Logistische und Multinomial-Logistische Regression inkorporiert, um die Bildungs- und Karrierewegsentscheidungen zu analysieren. Die Ergebnisse deuten darauf hin, dass Genderdifferenzen in der Selbstwahrnehmung weiterhin existieren. Die Konzentration auf die Selbstidentität könnte den politischen Entscheidungsträgern eine Erklärung bieten, warum finanzielle Anreize möglicherweise nicht ausreichen, um Personen für einige Mangelberufe zu gewinnen.

Contents

List of Tables

Introduction

The status of women in the socio-economic environment has been improving considerably in the last decades. Gender equality is a fundamental right and it is of high concern to the society, because it encourages economic development. Nevertheless, gender inequality varies from country to country and is still persistent in many aspects of life such as education and work.

The European picture reveals that women, who make the half of the population, are under-represented in leading positions in business and politics and still earn 16% less than men across the European Union (European Commission report on Equality, 2018). Why, despite the European initiative to promote gender equality, is the EU still behind from being gender-equal society? Is it because women are still perceived as less ambitious than men? Or because women self-select into the less professional careers? Becker (1985) provided one of the most prominent arguments to why women earn less: The earnings between men and women diverge because of the division of labour within families, while men specialize in career, women focus on housekeeping and child-rearing.

The differences in earnings and career opportunities may arise not only given the differences in the skills. The unexplained wage gap (the difference in earnings between men and women that have the same or similar skills to perform the same type of job), and gender gap (dis-proportional amount of men and women in certain job positions) could be caused by gender discrimination. Discrimination may arise even in the absence of prejudice when a member of particular minority group caries information about a person’s productivity and skills characteristics. Since 1960 due to the so called baby boom phenomena the labour force participation increased considerably, which resulted in the influx of working women. The promotion of gender equality in Europe dates back to Treaty of Rome of 1957. However, it was just recently that EU placed more attention to combat discrimination on a wider range of grounds of gender, race, age, belief etc. The adoption of EU Equality laws (2000) promote the inclusion and participation of disadvantaged groups. (European Commission. 2019). Today, the labour market is well regulated by the number of anti-discriminatory laws, which prohibit employers to discriminate.

Although, in the last decades women made it to the better representation in the leading positions, education and politics, the wage gap between men and women still persists and European organizations lack a structured pathway to make any improvement in female representation at the professional level and above in the next decades.(When Women Thrive, 2016). According to Global Report 2016 “When Women Thrive” women will make up 37% of those employees at professional level and above in 2025. Even though, the hire rates for women at the top of an organization are almost double those of men, mainly due to quotas, regulation or media pressure, this ‘quick fix’ is not working. As organizations fail to put into place supporting policies and practices, senior women are more likely to exit.

What is more, according to Eurostat (2018), the majority of women are employed in rather low-paid occupations. Statistics indicate that most common occupational choices for women are shops sales persons, personal care workers, school teachers and secretaries. However, the share of women that chooses occupations within STEM (Science, Technology, Engineering and Math) and ICT (Information and Communication Technology) is relatively low. According to the EU statistics, 42%, of all scientists and engineers in 2017, were women, however, men (83%) were highly overrepresented in the high and medium-high technology manufacturing. (Eurostat, 2019). What is more, ICT sector, that creates about 120000 jobs every year, employed 16.7% of women according to European Parliament. Many occupations require certain level of skills and education and some of the education decisions have to be made as early as in high school. Therefore, in order to understand these socio-economic outcomes it is imperative to study the education choices of an individual.

The traditional labour Economics textbooks suggest that the choice of education can be analyzed as an investment decision. This approach assumes that people acquire certain education level that maximizes their present value of lifetime earnings. However, the traditional economic analysis of education choice, fails to explain why skilled individuals choose low paying careers. Therefore, the educational policy incentives focusing on financial aspects are likely to be ineffective in influencing the educational paths/career choices individuals make.

An alternative approach is to integrate sociological factors to the education/career choice model. Akerlof and Kranton (2000) suggest that an individual's identity affects economic outcomes. The authors incorporate the psychology and sociology of identity (person's sense of self) into economic model of behaviour by adding the non-pecuniary payoff associated with the identity into utility function. They construct game-theoretic model to show how the identity affect interactions. According to the authors, there are many aspects related to identity driven behaviour. To name a few: (1) people have identity-based payoffs derived from their own actions; (2) people have identity-based payoffs derived from others' actions; (3) third parties can generate persistent changes in these payoffs; and (4) some people may choose their identity, but choice may be prescribed for others. The idea of Akerlof and Kranton (2000) could explain why highly skilled individuals choose low paid career.

Authors Humlum, Kleinjans and Nielsen (2007) apply the idea of non-pecuniary payoffs to study career choice and identity of young students in Denmark. They use factor analysis on a range of attitude questions and derive two factors related to identity, which they name "career" orientation and "social" orientation. Authors estimate logit and multinomial logit models to find the career and social factor effects on educational plans. They find that identity related factors are more relevant for women's choice between short and medium cycle educations, however, they do not matter for men´s choices.

The idea of this paper is to address the gender gap in certain occupations by analyzing education and occupation choices individuals make. The contribution of the thesis is that it incorporates socio-psychological factors into economic analysis to study whether non-pecuniary factors in terms of identity influence career choice.

This study uses the Swiss data set TREE (Transitions from education to employment). The TREE survey provides information about students' attitudes towards educational and career life. Based on the students' answers about their work attitudes and their non-cognitive abilities, we use factor analysis to construct students' identity measures. We identify four factors that represent identity and name them career, social, self-esteem and coping factors. We are interested in what factors influence individual choices regarding educations choices. In this thesis, we use logit and multinomial logit analysis to conduct the study. The traditional economic education choice models provide insufficient amount of information on why some male and female students, having the same skills or abilities choose different career paths. In addition to traditional cognitive measures, such as math and reading scores we incorporate the identity measures in order to study education and career choices of the students. This study provides some evidence in the context of Switzerland that gender differences are persistent in choosing different educational paths.

The structure of the thesis is as follows: The first section reviews the relevant literature and provides an overview of different approaches to study education/career choice. The second section provides overview of the education system in Switzerland. Section three presents the data. Section four presents methodology and model followed by the analysis in section five. Section six provides conclusion.

1. Previous Research

This chapter reviews the relevant literature and provides the overview of the previous research that lays the ground for this thesis. The departure point of this chapter is to provide the literature review of education selection that sheds light on the different career choices young people make. The focus of this paper is put on the supply side, namely, on individuals choosing and pursuing certain career paths. This chapter begins with the overview of the classical education selections models. Subsequently, socio-economic studies regarding education selection are presented. Moreover, it is also important to consider the demand side from employer's perspective, therefore, theory of discrimination will be presented. Finally, this chapter will introduce the alternative approach to study education selection. Having all these aspects in mind, one can gain a deeper understanding of the education selection.

1.1 Education and Occupation Selection Theories and Models

Discussions about the higher education policies may be hampered by the inability to predict the effects of the proposed policies on the student behaviour. There are number of theories that attempt to explain the student education selection behaviour. The university enrolments are the result of decision made by university administrators and by prospective students.

Authors Kohn, Manski and Mundel (1976) develop a theoretical and empirical model of student behaviour. They separate student decision problem into three stages: (1) for each available college, the choice of whether to commute or to live on campus; (2) the choice of the best college available, given residency decision; (3) the choice of whether to enrol at this best college or not at all. Enrolment follows if this best alternative is more attractive than the various possibilities other than college, such as technical education, the armed services, or immediate employment. The student's evaluation of a given college is based on the perceived costs and benefits of attending college. The authors recognize that the college is both an investment and a consumer good. The effective cost of attending an institution is determined by the institution itself, by government, by private groups through setting the tuition and living costs and through distribution of financial aid. Students vary in ability, location, income and family background so that colleges available, the costs and the benefits of a given college and the alternatives to going to college are different for each student. The authors assume that students self-select, a student will not apply to those colleges that s/he considers inferior, too expensive, or unlikely to admit her/him.

The authors attempt to estimate student's utility of going and not going to college. In order to estimate student's utility a number of assumptions must be borne in mind. The authors have information on the feasible college alternatives in their data set. The limitation is that the set of alternatives other than college faced by individual is not available. The authors use the conditional logit model to estimate the parameters. The model maximizes the likelihood of the observed choice.

The estimation showed that probability of going to college rises with the family income, as it goes from low to middle income strata, however, going to college falls sharply for those students from high-income families. Although, it would seem that students would prefer colleges with higher SAT averages, the results showed that a student would not want to attend college where average SAT score was too far from his own score. The preference for colleges with a wide range of programs was strong.

Authors Kohn et al. (1976) recognize demand for education is derived from expected lifetime earnings and its costs. Authors Willis and Rosen (1979) find that expected earnings are important determinants of the decision to attend college. However, the financial incentives alone do not provide a full picture to the college decision problem. Family background, tastes and 'ability bias' are important factors that influence the decision. The authors assert that 'Total variance of earnings among people of the same sex, race, education, and market experience is very large, and more than two-thirds of it is attributable to unobserved components or person-specific effects that probably persist over much of the life cycle'. (Willis and Rosen, 1979, p.S8).

The economic theory of education postulate that the schooling is pursued until the marginal private rate of return equals rate of interest. This leads to the recursive econometric model in which (one) schooling is related to the person´s ability and family background and (two) earnings are related to ´prior´ school decisions and ability. In the model proposed by Willis and Rosen, the earnings' gains attributable to the education do not appear explicitly in the schooling equation. Costs and benefits of alternative school-completion levels are assumed to be randomly distributed among people with different education financing capacities, tastes, expectations and talents.

There are many limitations to the estimation of selectivity models. The covariances of the unobservables are unrestricted. Also there may be negative covariance among talent components. As the authors note, for example, plumbers (high school graduates) may have very limited potential as highly schooled lawyers, while lawyers may have much lower potential as plumbers than those who actually end up choosing that kind of work. Another limitation, to the Willis and Rosen proposed model is that it is unknown whether a person chose college education because s/he was talented or because s/he was wealthy.

The authors find that expected gains in life earnings influence the decision to attend a college. It is also shown that financial constraint and tastes pose important effects. There is also effect of positive sorting or positive selection bias. This implies that people who stopped schooling after high school had better prospects of doing so while those who continued further education in college also had better prospects than an average member of sub-population.

Boskin (1974) has used the application of human capital theory to occupation choice. The individual worker weights the benefits, such as potential earnings, non-pecuniary payoffs and costs such as training and forgone earnings, etc. A worker will invest in changing the occupations only if the returns are sufficiently large and profitable in making the change of occupation. However, the author also assumes the market imperfection, meaning that resources for investing in one self is not equally available to all workers and that the wealth position of individual is important in making the decision. Thus, the decision on occupation choice will be based on the returns and costs. The author analyzes the choice of occupation among multiple alternatives using multinomial logit models. The probability of choosing a particular occupation j, is a function of the relative present values of potential post-investment lifetime earnings, E, trainings costs and foregone earnings relative to wealth, T/W_i, and the present value of expected income forgone due to unemployment, U, in alternative occupations:

$$P_{ij} = f\left(E_{i1} \dots, E_{ij}, E_{ij+1,} \dots E_{in}; U_{i1}, \dots, U_{ij}, \dots U_{in}; \frac{T_{i1}}{W_i}, \dots, \frac{T_{ij}}{W_i}, \dots, \frac{T_{in}}{W_i}\right).$$

In order to calculate the present value of earnings for a representative individual, the author uses 11 broad occupational classes:

- Professional/technical
- Farmer
- Manager
- Clerical
- Sales
- Craftsman
- Operative
- Private household
- Service
- Farm labourer
- Labourer

The findings are in line with the well-known labour market phenomena: Workers choose the occupations with the highest potential future earnings and those, where the costs of retraining in relation to net worth are the lowest. Moreover, the result show that white males tend to weight training costs and expected income foregone due to unemployment relative to expected full-time earnings much less heavily than other groups, such as white females and black males and females.

To sum up, a commonality among these theories is that they assume that individuals select the occupation that maximizes their utility function given some constraints set by the market and by their personal abilities. The education and career decision can be summarized as follows: Firstly, based on the market wages and ability endowments, an individual calculates expected lifetime income for all relevant occupations. Secondly, an individual compares the utilities of going to college or not going. And finally, individual selects the career path which offers the greatest utility.

The economic theory of human capital explains why individuals choose to invest in their human capital via education and how these choices affect the future earnings. Economic theory suggests why some young people choose to pursue more higher education and why some choose to drop out early. People who invest in schooling are willing to give up some earnings today for future higher earnings. The trade-off of forgone earnings today and future earnings influences persons' educational attainment. Thus, most of the economic theories of education and career choice mainly focus on the effect of wages. The economic theory does not deny the importance of the abilities or preferences in the career decision. However, it is difficult to account for sociological and psychological phenomena in the economic career choice models. What economic theory does it just takes these effects as given and focuses on wage effects.

The educational attainment is important for employability. The lower educated are the first to be fired and stand at the back of the job queue when there are vacancies to be filled in. Wolbers (2000) finds that after several months of unemployment, the chances of individuals with only primary-level education finding work again in the following month are slightly more than 2 percent, while for those with a secondary-school diploma or graduates of higher vocational education chances are more than 4 percent and for university graduates about 6 percent respectively. Thus, education is associated with higher employment rates and higher earnings.

1.2 Socio-economic Studies of Career Choice

Career relevant education choices are not made in the social vacuum. Therefore, it is important to recognize the social influences as well as economical reasoning in the education and later career choices. This chapter provides the theoretical framework of career choices from different perspectives: economic and social. Later attempt will be made to combine these social perspectives in economic career choice model to gain further understanding of individual career choice.

Social cognitive career theory (SCCT) was developed to help explain the interplay among person variables (e.g. Self-efficacy, outcome expectations, goals) and contextual variables (gender, culture, support systems, and barriers) that influence three important phases of career development: (1) the formation of vocational and academic interests, (2) selection and pursuit, performance and (3) persistence in educational and occupational endeavours (Lent and Brown, 2001).

Factors influencing people's decision making regarding college major might be both internal (self-efficacy and outcome expectation, coping efficacy) and external (social contextual barriers and supports). An internal factor self-efficacy refers to the perceived capability to perform particular behaviours for success within certain domain. Whereas coping efficacy, reflects one's perceived capability to negotiate particular situational features that complicate the performance. For example, student may believe that s/he has strong math capabilities (task self-efficacy) yet lack confidence at withstanding negative peer pressure linked to pursuing math related major.

The authors Lent and Brown (2001) focus mainly on the external factors and explore whether environmental supports and barriers moderate interest-choice relations. In other words, whether contextual supports or barriers influence the process where people translate their career interests into goals and their goals to actions. Authors find that contextual variables to choice outcomes had few significant relations between barrier perceptions and career outcomes. However, they found that barrier perceptions are related to coping self-efficacy. These findings suggest that influences of barriers and support perceptions on the choice making process have indirect effects on individuals.

Siann and Callaghan (2001) in their paper “Choices and Barriers: factors influencing women’s choice of higher education in science, engineering and technology“ suggest that women are deterred by the nature of scientific enquiry. Since scientific enquiry has until very recently been mainly conducted by men, the most fundamental aspects of scientific though have been pervaded by masculine perspectives derived from masculine experiences. Keller (1983) argues that children identify scientific though with masculinity because it is so deeply embedded in the culture. Many children grow up not only expecting scientists to be men, but also perceiving scientists as more “masculine” than other male professionals, than, for example, those in the arts. (Keller, E.F. in Harding and Hintikka, 1983).

There are number of studies that go beyond focus on human capital and look at the impact of social factors such personality traits on differences on career chances. The systematic way to organize the personality traits can be traced back to McDougall (1932) as he suggested that although personalities are infinitely various and complex, they can be broadly analyzed into five distinguishable but inseparable factors such as intellect, character, temperament, disposition and temper. Over the years researchers have investigated the personality measures and the concept of the Big Five has emerged, which is widely adapted today. It is widely agreed that first dimension is Extraversion/ Introversion. The second generally agreed dimension is Emotional stability, or Neuroticism. The third is Agreeableness. The fourth is Conscientiousness. The fifth dimension has been more difficult to identify, but is generally interpreted as Intellect. (Barrick and Mount, 1991).

The trait theory has been used in many studies. The field of leadership research focuses on personality traits that differentiate the leaders from the other employees. Fietze, Holst and Tobsch (2011) investigate whether personality can explain the gender career gap in Germany. They focus on a Big Five personality traits and also willingness to take risks in order to estimate the likelihood of being in a leadership position. They show evidence that women in leadership positions differ more from their female colleagues who are not in leadership position than men. The results show that probability of being in a leadership position is greater for employees who are emotionally more stable, more open to experience, less agreeable and more conscientious. When assessed separately, it is shown that women can increase their probability of being in the leadership position through less agreeableness, while men can increase their probability of being in the leadership position through conscientiousness and emotional stability. For both sexes, willingness to take risk in the career paths has the largest impact.

Spearman’s (1904) research on human intelligence led him to develop a theory of factor analysis. In his work “General intelligence” he examined students hearing, touch and sight senses as these senses were regarded as best mental activities in measuring intelligence. Spearman measured correlations between mental activities and observed that there must be some common faculty that is driving cognitive abilities. “Whenever branches of intellectual activity are at all dissimilar, then their correlations with one another appear wholly due to their being all variously saturated with some common fundamental Functions (or group of Functions)" (Spearman, 1904, p.273). Today, Spearman's work on general intelligence is known as the g-theory.

Raymond Cattell, a psychologist known for his research on the trait theory, further developed Spearman's factor analytic methods and applied them to the psychological research. Cattell was a proponent of using exact mathematical methods instead of vague terms in defining personality traits. Given the common position of trait theory that personality can be described in terms of discrete if not independent traits, Cattell claimed that in the empirical statistical view that a trait exists where the inter-correlations of trait elements form a cluster of high values, when there is an 'operational unity' (Cattell 1945). He used factor analysis in establishing trait unities, which are calculations of correlations that look for common variations in trait variables in the field of individual differences. Cattell further developed Spearman's general ability factor (g), expanding the general ability factor. He developed the theory of "fluid and crystallized" general abilities, stating that general ability factor measured by intelligence test would be found to be not one but two factors. "Crystallized" intelligence loads highly on those cognitive performances in which skilled judgement habits have become crystallized (hence the name) as a result of earlier learning application of prior ability (for example, verbal and numerical abilities). Also crystallized intelligence is partly a product of motivational and personality history, that has more significant associating with the personality factors. While, "fluid general ability" shows more in tests requiring adaptation to new situations. (Cattell 1963).

1.3 Discrimination in Labour Market

The differences in earnings and career opportunities may arise not only given the differences in the skills. Other characteristics of labour force such as race, gender, national origin, etc. may determine wage dispersion. These differences are often attributed to the labour market discrimination. The discrimination arises when employees belonging to certain minority groups are treated differently than members of majority groups of equal productivity.

One of the most prominent works to lay the ground for the study of discrimination was conducted by Gary S. Becker (1957). In the book 'The Economics of Discrimination', the author develops a concept called 'taste discrimination'. The concept of 'taste' translates into the prejudice in the labour market. To put it formally, if an employer is prejudiced against some minority group members, such as blacks, the employer gets disutility from hiring black workers. Thus, the employer's distaste for black workers imposes the employer to act as if it costs more to hire black workers as it actually is. To put it in economic representation, w_b is the wage rate of black employee, if the employer is prejudiced, then it costs $w_b(1+d)$ to hire a black person, where d is a positive number, which refers to the discrimination coefficient. However, not only employers may have taste for discrimination, this term may apply to other economic actors: employees and customers may also act in a discriminatory ways.

The taste for discrimination reflects a subjective view, however, discrimination may arise even in the absence of prejudice when a member of a particular minority group caries information about a person's productivity and skills characteristics. This refers to the statistical discrimination.

1.4 Statistical Theory of Labour Discrimination

The Statistical Theories of Discrimination try to explain the phenomenon of discrimination in the light of employer uncertainty about the productivity of racial (or gender) groups of workers, particularly in hiring decisions. Phelps (1972), Aigner and Cain (1977) highly contributed to statistical discrimination theories.

Phelps (1972) suggests that the employer who seeks to maximize expected profit will discriminate against blacks or women if s/he believes them to be less qualified, reliable, long-term, etc. on the average than whites and men, respectively. Given uncertainty in hiring, employer may form a priori belief about profitability of white over black (male over female) which is based on the employer's previous statistical experience with the different group's members. Thus, members of favourable group might be chosen over the members from less favourable group.

In an economic model of statistical discrimination proposed by Phelps, employer can observe applicants performance test score, which may be used as a measure of the applicant's qualification plus an error term. No other additional information is available to the employer except for the test score and applicant's skin colour. Author adjusts the model accordingly and includes a race related term, so that the test data can be used in relation to race (sex) factor to predict the degree of qualification. Consequently, if employer believes that blacks have some social disadvantage, then it is expected to have lower prediction of qualification for blacks than whites having equal test scores.

In a further attempt to explain the statistical discrimination Aigner and Cain (1977) argue, that Phelps's model despite laying the ground for statistical discrimination, fails to provide explanation for most discrimination scenarios. Aigner and Cain (1977) modify the model proposed by Phelps such that the expected value of qualification, which depends on the observable test score, can be expressed in terms of group effect and individual effect. The later model presents the idea of statistical discrimination in a much more intuitive way than Phelps's model, because it is directly visible, that the expected qualification is weighted on the average group performance, that members belongs to (hence the name statistical discrimination). In contrast to Phelps (1972), Aigner and Cain (1977) do not assume that the variance of white qualification is less than that of blacks, because discrimination defines the difference in pay that is not related to productivity. Instead, authors incorporate the risk factor that an employer may face when making hiring decision, which may explain why blacks receive lower wage than whites on average for the same expected ability.

These theories give insights to possible race discrimination in labour market, the same principle could also be applied for cases of gender discrimination. However, this is only one side of a coin and there is much more to discrimination that influences the wage differences between races and in the case of this work's interest: the gender remuneration differences.

The workforce at the end of the 20th century and beginning of 21st century is different from workforce that prevailed before. Since 1960 due to the so called baby boom phenomena the labour force participation increased considerably, which resulted in the influx of working women. At the same time first anti-discriminatory laws started entering into force. The Title VII of Civil Act of 1964 in US prohibit employment discrimination on the basis of race, colour, religion, sex, and national origin. Today, the labour market is well regulated by the numerous anti-discriminatory laws, which prohibit employers to discriminate, however, with the fast changing work environment of the 21st century there are other challenges to come.

1.5 Education Selection and Identity

As previous studies show, there are other important factors besides general abilities influencing people choices. An alternative approach to traditional one is to integrate sociological factors to the education/career choice model. Akerlof and Kranton (2000) suggest that individual's identity affects economic outcomes. Previous economic analyses of labour supply and schooling have not considered these aspects.

The authors propose the following utility function:

$$\mathbf{U}_{\boldsymbol{j}} = \mathbf{U}_{\mathbf{j}}(\mathbf{a}_{\mathbf{j}}, \mathbf{a}_{-\boldsymbol{j}}, \mathbf{I}_{\mathbf{j}})$$

$$\boldsymbol{I}_{\boldsymbol{j}} = \boldsymbol{I}_{\boldsymbol{j}}(\mathbf{a}_{\mathbf{j}}, \mathbf{a}_{-\boldsymbol{j}}, \mathbf{c}_{\mathbf{j}}, \boldsymbol{\epsilon}_{\boldsymbol{j}}, \boldsymbol{P})$$

Utility depends on j's identity or self-image I_j, as well as on the usual vectors of j's actions, a_j, and others' actions, a_{-j}. Since a and a_{-j} determine j's consumption of goods and services, these arguments and U_j(-) are sufficient to capture the standard economics of own actions and externalities.

Where person's j identity depends on j's assigned social categories c_j. Identity further depends on the extent to which j's own given characteristics ϵ_j match the ideal of j's assigned category, indicated by the prescriptions P.

To summarize, identity is bound to social categories. It depends on how people perceive themselves and how others perceive them. For example, female trial lawyer, male nurse, woman marine all conjure contradictions. Why? Because trial lawyers are viewed as masculine, nurses as feminine, and a marine as the ultimate man. People in these occupations but of the opposite sex often have ambiguous feelings about their work. In terms of the utility function defined above, an individual's actions do not correspond to gender prescriptions of behaviour.

The authors Humlum et al., (2012) set up a model of identity and career choice. In their model the career choice is equivalent to the education choice. When identity is introduced into the utility function, an individual takes into account that his/her choice of effort e_i put into educational attainment affects his/her self-imagine I_i. Thus, the utility function becomes:

$$\boldsymbol{U_i = U_i\big(w_i(e_i, \epsilon_i), e_i, I_i(e_i, c_i, \epsilon_i, P)\big)},$$

Where $w_i(e_i, \epsilon_i)$ is the income of individual i given career choice e_i and characteristics ϵ_i ,c_i assigned social categories and P prescriptions.

In order to test this theoretical model the authors construct the measures of identity. They use data for Denmark from PISA and a follow up survey (PISA FUS). In these surveys participants are asked questions for plans of education. Authors use this information to derive measures for planned education: one for planned level and other for the field. They use factor analysis to derive two orthogonal factors capturing an individual's identity: one factor is labelled as "Career orientation" and another factor that is labelled "Social orientation". These factors load heavily on questions related to general attitudes about career and society, which are closely related to person's social identity and work life. The other explanatory variables used in their empirical analysis include ability measures, number of individual characteristics, family characteristics, such as parental socioeconomic status and information on birth order. These variables may be part of prescribed characteristics or they may reflect tastes for education.

Authors estimate logit and multinomial logit models to find the effect of the career factor and social factor on educational pans. The "Career orientation" factor and "Social orientation" factor scores are predicted for each individual, which are used as explanatory variables. To quantify the effects of the factors, authors estimate conditional logit models which provide dollar equivalent value of the factor scores. In order to obtain these values, they include counterfactual predicted annual wage incomes for each possible education plan as an additional explanatory variable and then compare the effect of factor scores on the career choice to that of the predicted wage income.

Findings show that for women a higher career factor increases level of education, while a higher social factor decreases it. There are no effects found for men's behaviour. When it comes to field of education, they find that one standard deviation increase in the career factor moves 7% of the youth from education and humanities to business, law and social sciences, whereas one standard deviation increase in the social factor moves 9% of the youth away from business, law and social sciences into other fields such as health sciences.

Heckman, Stixrud and Urzua (2006) establish that cognitive and non-cognitive skills explain a variety of labour market and behavioural outcomes. They use factor analysis to measure cognitive and non-cognitive skills. The assumption that one latent factor captures cognitive ability is traditional in the literature. The tests that measure the skill in following areas: general science, arithmetic reasoning, word knowledge, mathematical knowledge and coding speed comprise the cognitive factor. However, the assumption of latent factor of non-cognitive ability is less traditional. Authors measure the non-cognitive skills with the help of attitude tests: Rotter Internal-external Locus of Control and Rosenberg Self-Esteem. They use the NLSY79 data for their analysis.

Authors construct different models using the factors and the observed variables. They assume that the latent cognitive and non-cognitive abilities, denoted by f^C and f^N, are mutually independent and that they determine the individual's wage, schooling, employment, work experience and occupational decisions.

The model fits the data of NLSY79 on different outcomes such as schooling level and occupational choice. The data analysis is carried out separately for males and females. The educational choice model is estimated using multinomial probit model, which considers six different categories: high school dropouts, GED recipients, high school graduates, some college but no degree, 2-year college graduates, and 4-year-college graduates.

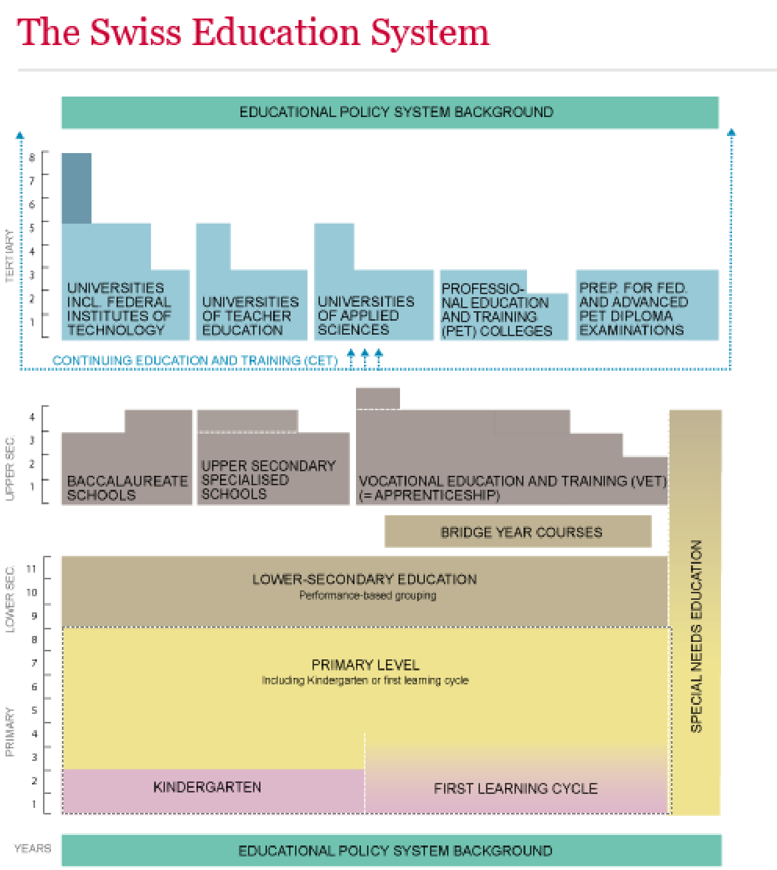
The occupation model is estimated using probit model. The dependent variable takes a value of 1 if the individual report a white collar type of occupation, or 0 if blue collar occupation. For both genders, cognitive and non-cognitive abilities are important determinants of the choice of white- versus blue-collar occupations.

The loadings on both cognitive and non-cognitive factors are statistically significant in most equations. Both factors are required to produce a model that passes goodness-of-fit tests. The estimated distributions of the factors are highly non-normal.

The authors find that latent non-cognitive skills raise wages through direct effects on productivity, as well as through indirect effects on schooling and work experience. Their evidence is consistent with an emerging body of literature that finds that "psychic costs" (which may be determined by non-cognitive traits) explain why many adolescents who would appear to financially benefit from schooling do not pursue it.

2. Education System in Switzerland

The analysis of education/career choices in this paper is based on the dataset for Switzerland. In order to conduct our analysis, it is important to have a general overview of the education system in Switzerland, which is provided in this section.



Picture 1. (2016). The Swiss Education System. (Swiss education)

Every child in Switzerland has to complete the compulsory education. The total period of compulsory education amounts to eleven years. Primary level – including two years of kindergarten or a first learning cycle – comprises eight years. Lower secondary level takes three years. Children in these grades tend to be between 12 and 15. There is no nationwide exam at the end of ninth grade – the final year – so students receive no graduation certificate.

After completing the compulsory lower secondary education students may transfer to the upper secondary education. More than 90% of young people complete an upper secondary level programme. The adolescents complete upper secondary level at the age of 18/19 and receive a corresponding certificate.

Upper secondary education can be subdivided into general education programmes, and vocational education and training (VET) programmes.

The majority of adolescents enrol in vocational education and training (VET) after lower secondary level. There are VET programmes for some 250 different professions. In Switzerland, many professional qualifications are obtained in upper secondary level, while in other countries the same qualifications are obtained in tertiary level education. The Swiss system, therefore, differs from most foreign systems of vocational and professional education and training. VET is predominantly based on a dual system: practical training (apprenticeship) of three to four days at a training company is supplemented by theoretical classes (vocational and general educational subjects) on one to two days at the VET school. (Swiss education, 2019). More detailed description of the Swiss education system is provided in the appendix.

3. The Data

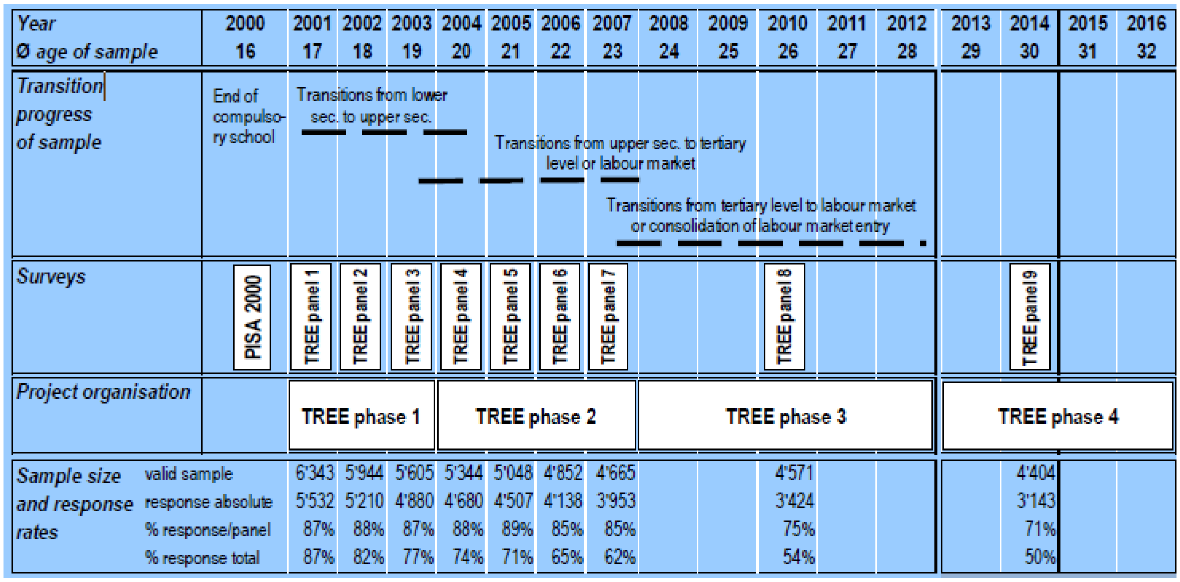
This study uses the Swiss panel study TREE (Transitions from Education to Employment), which is a social science data infrastructure. The TREE study is mainly funded by the Swiss National Science Foundation (SNF) and located at the University of Berne. The TREE is a multi-cohort survey which follows two large samples (>6,000 respondents) on their pathways through post-compulsory education and employment well into young adulthood.

TREE1 was launched in the year 2000 on the basis of the Swiss PISA (Programme for International Student Assessment) sample. It contains nine waves of surveys through year 2001 until 2014. Today, this sample has reached an average age exceeding 30, and the data collected so far cover an observation period of approximately fifteen years. “Transition pathways to employment are becoming increasingly complex. To understand young people’s decisions and options, and to take them into account in policy decisions, appropriate analytical instruments are needed.” (OECD 1999: 53).

The TREE dataset enables to check for relationships between socio-economic and socio-cultural factors, skills and competencies, personality traits and career ambitions as well as features of the school environments at the end of compulsory education and the actual education and employment careers pursued.

The work presented in this thesis combines PISA data and TREE data set to have a wide range of the relevant variables. The PISA survey contains measures of language, math and science skills. The TREE data contains information on personality traits and career ambitions that are relevant for this study about education/occupation choice.

The work presented in this thesis focuses on the 2002 and 2007 waves of survey. The 2002 wave shows the first tendencies of choosing different educational paths, while in 2007 wave most of the participants have attained the education and are in the employment.



Picture 2. Source: TREE

In the PISA-TREE survey, wave 2002, there are 6343 participants out of which females account to 3440, or 54.22%. The response rate of the survey in 2002 is 88%, however, the response rates of separate questions are often lower.

The current status of education picture reveals that most of the 2002 wave participants chose vocational training, which is 573 out of 1035. 428 are in gymnasium and 34 do not continue with any education. 50% out of all females choose gymnasium and 47% choose vocational path, while 67% of all males choose vocational path and 29% choose gymnasium.

The 2002 wave includes questions that describe career values, where respondents indicate what they value in the future job. The responses to the question 'It is important to have a job where I can help other people' 43% percent indicate that it is rather important to them, and 35% indicate that is very important. However, when we compare female and male response rates, we see that females put more weight on the importance of helping other people in their jobs. (43% females indicate that it is very important to them, while this is very important to 24% of males). To most of the male and female respondents, it is very important to have a job that has a meaning. When asked 'It is important to earn a lot of money both female and male regard it rather important. However, 41% of males consider having a good wage very important, while 27% females consider it very important. (See appendix, tables from 'steps' about values).

PISA measure on reading literacy reports the reading proficiency of 15 years old as they approach the end of compulsory education. Reading literacy measure represents major aspects, which are tested on retrieving information from variety of reading materials, interpreting what is read, reflecting upon, and evaluating what is read. Performing these tasks involves wide variety of cognitive abilities. PISA survey measure on reading literacy shows that majority of students, 33% score medium high, 25% score high and 13% percent score very high. Females tend to score higher on reading literacy. 33% of females score high as compared to 24% of males that score high, and 20% of females score very high as compared to 14% of males that score very high in reading literacy (Table 134 in Appendix).

In PISA survey, students are asked about their perceived occupation at the age of 30. Majority indicate that they still do not know. 579 out of 5785 indicate that they want to be Physical, mathematical and engineering science professionals and out of these 579, 345 indicate that they want to be Computing professionals. 357 indicate that they want to work as Life science and health professionals. 351 indicate that they want to be teaching professionals. 586 other professionals such as business professionals, legal professionals. 243Physical and engineering science associate professionals. 376 service workers and shop and market sales workers. 118 elementary occupations as street vendors, domestic helpers, construction workers etc.

The highest education attained by respondents of 2007 wave reveal that 1521 out of 2931 have finished or pursue vocational/professional education, 1381 have attained or are attaining university education, and 29 pursue no education at all. The most typical duration of studies is 3 years. Most of the respondents (644 out of 1877) that are still in education are in their 3rd year of studies at the date of survey wave 2007.

4. Methodology and Model

This study will follow the Akerlof and Kranton (2000) idea that there is a non-pecuniary payoff to the career choice through the influence of person's identity. In order to conduct the empirical analysis, the study approach of Humlum et.al (2012) will be adopted in this study. The empirical analysis will be conducted with statistical software STATA using PISA database, which provides data for OECD countries and Swiss data set TREE (Transitions from Education to Employment).The TREE data set contains all variables assessed in PISA 2000 for the entire TREE sample. PISA 2000 is the baseline survey of the cohort under observation. These surveys contain responses of the students, school principals and parents. The PISA part of the survey contains data on the student abilities, background, while the follow up surveys (TREE) contain information on the educational and employment status, on students attitudes towards educational and career life and coping abilities.

The aim of the study is to investigate whether student's identity, which is separated from student's ability, is associated with the choice of education. As Burke and Reitzes (1981) suggest, identities are meanings one attributes to oneself in a role (and that others attribute to one). Identities are social products. The individuals learn the meanings of self through responses to one's own actions. Also individuals are motivated to formulate plans and achieve levels of performance or activity that reinforce, support and confirm their identities.

4.1 Measuring Identity: Factor Analysis

Some variables such as ‘identity’ may not be directly observable. The latent variable models provide an important tool for analysis of multivariate data. According to Bartholomew (2011) the latent variables or factors help to quantify some hypothetical phenomena. For example, business confidence is often spoken as if it was a real variable. Yet business confidence is a concept, which may be regarded as a complex of beliefs and attitudes. Latent variables provide the framework for constructing such hypothetical concepts. Factor analysis operates on the notion that measurable and observable variables can be reduced to fewer latent variables that share a common variance and are unobservable. Thus one of the reasons of use of latent variables is to reduce dimensionality. Large-scale social surveys provide a lot of information and sometimes there is a need to summarize the data. But with many variables it is difficult to see the interrelationships. Latent variables or factors provide a way of condensing the variables into a smaller number of indices with little loss of information. Therefore, in order to classify individuals into different social categories/identities, the factor analysis will be used.

Factor analysis has its origins in its use in psychology and education. It was Galton, who laid the foundations of factorial study in the 19^{th} and early 20^{th} centuries. (Child, 2006). Galton’s idea lies on the form of g - general ability, formulated using factors solutions. Then Spearman’s development of the Two-factor theory in 1904 on intelligence and human ability using mathematical model laid the grounds increasing work on the theories and mathematical principles of factor analysis. Today factor analysis is used in many fields such as behavioural sciences, medicine and economics.

There are different techniques used in the factors analysis. One of the widely used techniques is Exploratory Factor Analysis (EFA), which will be employed in this study. EFA attempts to uncover complex patterns by exploring the dataset. This technique is used when the researcher wants to discover the number of factors. A basic idea of EFA is that there *m* common latent factors to be discovered in the dataset with the goal of finding the smallest number of common factors that will account for the correlations. (Yong and Pearce, 2013).

The unmeasured variables can be revealed by a particular structure in the observed correlation matrix. The general way of uncovering the structure is done by factor extraction and factor rotation. First factors are extracted. One of the most common ways of factor extraction is called ‘Principal Factors’. In this way factors are extracted from the part of the correlation matrix that has the proportion of the variance of each variable that it has in common with the other variables. This proportion is called communality of the variable. The goal of principle factors is to extract factors in such a way as to explain the maximum amount of variance. (Harman Harry, 1976). Since in this thesis we use exploratory factor analysis, we do not predetermine how many factors to extract. As a general rule, the number of factor is extracted based on a number of eigenvalues greater than one.

After factors have been extracted, their weights are not generally interpretable, therefore, we need to use factor rotation. The most common rotation method is called Kaiser’s Varimax procedure, which maximizes the variance of the squared loadings within each column.(Harman Harry, 1976). Once factors are rotated, one can start interpreting the factors. Researcher examines the variables and their loadings on the factors and tries to find patterns. It is common to regard factor loadings greater than 0.3 as 'significant'. Bartholomew (2011). Generally, at least two or three variables must load on a factor. After the examination of factor loadings, the researcher creates the factor labels given the variable interpretation that load on the respective factors.

4.2 Qualitative Choice Analysis

Once the factors are identified, they can be used to conduct multinomial logit analysis to investigate the effects of identity on the choice of education. The logit models are known to be used for analysis of qualitative choice behaviour such as choice of college, choice of occupation, mode of transportation, etc. The alternative could be a probit model. However, because of the need to evaluate multiple integrals of the normal distribution, the probit model has found rather limited use in this setting. The logit model, in contrast, has been widely used in many fields, including economics, market research, politics, finance, and transportation engineering. (Greene, 1951, p.761).

One of the implications to the multinomial logit models is the assumption of the independence of irrelevant alternatives (IIA). IIA mean that the person's choice between two alternative outcomes is unaffected by what other choices are available. However, this assumption can be unrealistic as illustrated by the 'red bus' and 'blue bus' example by McFadden (1973). As McFadden has noted the, multinomial logit models should be used when the outcomes can plausibly be assumed to be distinct and weighted independently in the eyes of each decision maker. We assume that all occupations have a positive selection probability for each individual and that the odds that a particular occupation will be chosen over another are independent of the presence of other possible occupations. Therefore, in this study we apply multinomial logit model.

The idea of the multinomial logit model is that individuals choose among more than two choices, making the choice that provides the greatest utility. The observed outcome is the count of the number of occurrences, therefore the model is focused on the discrete outcomes in terms of probabilities attached to these outcomes. Suppose individual i is faced with J choices and utility of choice j is:

$$U_{ij} = z'_{ij}\theta + \varepsilon_{ij}.$$

Where, variable z_{ij}, includes aspects specific to the individual and to the choices.

If the individual, choose choice j, then it is assumed that the U_{ij} is the maximum among the J utilities and, thus, the model is driven by probability that choice j is made:

$\mathrm{Prob}(U_{ij} > U_{ik})$ for all other $k \neq j$.

4.3 Theoretical Framework

This study adapts the utility function suggested by Humlum et.al (2012)

$$U_i = U_i\big(w_i(e_i, \epsilon_i), e_i, I_i(e_i, c_i, \epsilon_i, P)\big)$$

Where e_i represents career choice, which can also be thought of as extent of effort exerted for example the level of education, but also different types of effort for example field of education. I_i represents identity, c_i assigned social categories and P prescriptions as in Humlum et al (2012).

The variables, reflecting the personal characteristics as coping abilities and attitudes towards job, such as enjoyment working in the teams, having secure job position, to have position with career opportunities, etc. will be used to construct the factors. Humlum et al. (2012) identify two factors: a career factor and a social factor which reflect a career oriented category and socially oriented category respectively. A career factor is interpreted based on how heavily it loads on the questions regarding the statements about importance of career and work, while social factor is interpreted as it load heavily on questions regarding statements about importance of cooperation, social issues. It is expected that, those individuals that have higher career factor score, would choose the "Business, law and social sciences", while those individuals having higher social factor would choose "Health sciences" or other humanitarian sciences.

In order to analyze planned level of education, we use logit and multinomial logit analysis.

a) We group education level/type into vocational training, gymnasium or no education in the sample of the 2002 in order to study what choices regarding education individuals make.

b) We repeat the same procedure in the sample of 2007, where we group education level/type into vocational, university and no education.

We set up a model that applies to each case:

1)

$$Education\ Level_i = \beta_i X_i + \alpha_i F_i + e_i\ , for\ each\ individual\ i,$$

Where, EduLevel is dependent variable, which represents choice of education level:

a) 0=Vocational, 1=Gymnasium.
b) 0=Vocational, 1= University.

And where, **X** and **F** are characteristics: X's include parental education, family wealth, individuals' cognitive ability, gender, age, agglomeration, and Fs are latent variable measures, that represent coping, self-esteem, career and social factors. For simplicity of notation we define characteristics as $(F_i+X_i)=w_i$.

Thus the probability of choosing certain education level is:

2) $Prob(EduLevel_i = j \mid w_i) = P_{ij} = \frac{exp(w_i^{'} \alpha_j)}{1+exp(w_i^{'} \alpha_j)}$, $j = 0,1.$

The estimated equations provide a set of probabilities for the J+1 choices, for the decision maker with characteristics w_i.

Next, we set up a model for the choice of occupation:

3) $Job\ Field_i = \beta_i X_i + \alpha_i F_i + e_i\ , for\ each\ individual\ i,$

Where 'JobField' is dependent variable, which represents occupations choice: 0=Physical, Math, Engineering professionals and associates, 1=Teaching professionals and associates, 2=Social sciences, Business, Law professionals and associates, 3=Life and Health science professionals and associates,4=Clerks/Office clerks, 5= Service workers, 6= Agriculture, 7= Craft, Plant and Machine operators.
And where, **X** and **F** are characteristics: Xs include parental education, family wealth, individuals' cognitive ability, gender, age, agglomeration, and Fs are latent variable measures, that represent coping, self-esteem, career and social factors.

Probability of choosing certain occupation:

4) $Prob(JobField_i = j \mid w_i) = P_{ij} = \frac{exp(w_i^{'} \alpha_j)}{\sum_{j=0}^{7} exp(w_i^{'} \alpha_j)}$, $j = 0,1,2, 3, 4, 5, 6, 7.$

The model implies that we compute J log-odds:

5) $ln\left[\frac{P_{ij}}{P_{ik}}\right]$,

It is useful that the odds ratio P_{ij}/P_{ik}, does not depend on the other choices, however, it is not so convenient for the interpretation of the coefficients. Therefore, we compute marginal effects of the characteristics on the probabilities of a specific choice:

6) $\delta_{ij} = \frac{\partial P_{ij}}{\partial w_i}$,

4.4 The Models

In the first specification, we only add gender dummy variable. Following we add more explanatory variables. In the second specification we add ability measures, in the third specification we add identity factors and in the fourth specification we add family tertiary education dummy variables, agglomeration dummy variable, family wealth and age. The correlation table and summary statistics can be found in the appendix. (Table 36 and 37). We apply models 1-4 in the analysis of the 2002 and the 2007 survey waves, regarding the education field choices. The models 5-8 are applied to the analysis of the 2007 survey wave, regarding the job field choices.

Model 1:

$$Education\ Level_i = \beta_i Female_i + e_i\ , for\ each\ individual\ i,$$

Model 2:

$$Education\ Level_i = \beta_{1i} Female_i + \beta_{2i} Reading\ Score_i + \beta_{3i} Math\ Score_i + e_i\ , for\ each\ individual\ i,$$

Model 3:

$$Education\ Level_i = \beta_{1i} Female_i + \beta_{2i} Reading\ Score_i + \beta_{3i} Math\ Score_i + \alpha_{1i} Social\ Factor_i + \alpha_{2i} Career\ factor_i + \alpha_{3i} Coping\ Factor_i + \alpha_{4i} Esteem\ Factor_i + e_i\ , for\ each\ individual\ i,$$

Model 4:

$$Education\ Level_i = \beta_{1i} Female_i + \beta_{2i} Reading\ Score_i + \beta_{3i} Math\ Score_i + \beta_{4i} Family\ Wealth_i + \beta_{5i} Mother\ Edu_i + \beta_{6i} Father\ Edu_i + \beta_{7i}\ Agglomeration_i + \beta_{8i} Age_i + \alpha_{1i} Social\ Factor_i + \alpha_{2i} Career\ factor_i + \alpha_{3i} Coping\ Factor_i + \alpha_{4i} Esteem\ Factor_i + e_i\ , for\ each\ individual\ i,$$

Model 5:

$$Job\ Field_i = \beta_i Female_i + e_i\ , for\ each\ individual\ i,$$

Model 6:

$$Job\ Field_i = \beta_{1i}Female_i + \beta_{2i}Reading\ Score_i + \beta_{3i}Math\ Score_i + e_i\ , for\ each\ individual\ i,$$

Model 7:

$$Job\ Field_i = \beta_{1i}Female_i + \beta_{2i}Reading\ Score_i + \beta_{3i}Math\ Score_i + \alpha_{1i}Social\ Factor_i + \alpha_{2i}Career\ factor_i + \alpha_{3i}Coping\ Factor_i + \alpha_{4i}Esteem\ Factor_i\ + e_i\ , for\ each\ individual\ i,$$

Model 8:

$$Job\ Field_i = \beta_{1i}Female_i + \beta_{2i}Reading\ Score_i + \beta_{3i}Math\ Score_i + \beta_{4i}Family\ Wealth_i + \beta_{5i}Mother\ Edu_i + \beta_{6i}Father\ Edu_i + \beta_{7i}\ Agglomeration_i + \beta_{8i}Age_i + \alpha_{1i}Social\ Factor_i + \alpha_{2i}Career\ factor_i + \alpha_{3i}Coping\ Factor_i + \alpha_{4i}Esteem\ Factor_i\ + e_i\ , for\ each\ individual\ i,$$

4.5 Hypotheses

Hypothesis I: Non-cognitive abilities in terms of personality and identity matter for education/occupation choice.

Hypothesis II: 'Career orientation' would lead individuals choose longer cycle studies and pursue careers in high paying occupations, such as in STEM occupations, while 'social orientation' would lead individuals to choose career in humanities, 'health science occupations'.

5. Analysis

5.1 Constructing Factors. Survey Wave 2002

In this chapter the analysis will be presented: it starts with factor analysis. In order to conduct factor analysis few things should be kept in mind. First, it will be looked at if the data is suitable for factor analysis. Second, it will be discussed how factors will be extracted and what criteria will assist in factor extraction. Next, the rotation method will be chosen and finally the interpretation and labelling of factors will follow. Once the factors are identified the multinomial logit analysis on education level and occupation field will be conducted. The analysis is performed on the two survey waves: first, we conduct analysis on survey wave of year 2002 and, second, of 2007.

For factor analysis we use a set of questions regarding respondents´ values and attitudes towards work. These questions have the ordinal response rates on a scale from 1 to 4, where ‘1’ is ‘totally subordinate’, ‘2’ is ‘rather subordinate’, ‘3’ is ‘rather important’, ‘4’ is ‘very important’. Missing values were excluded. The total number of questions included in the factor analysis is 27.

We need to consider some requirements and assumptions about factor analysis. Some authors suggest that there should be normality within data. (Yong and Pearce, 2013). Others suggest that normality is desirable but only required when factors are extracted by maximum likelihood. Next requirement is sample homogeneity. In our case the sample is homogenous because it is drawn from the same age population, from the same location. Furthermore, sample size should be adequate. Fabringer and Wegener (2011) suggest that under optimal condition sample of 100 can be adequate and there should be 3 to 5 measured variables loading on factor. In our case, we have from 961 to 1029 observations for each question used in factor analysis. To find factors we use correlations between the variables. Therefore, we need to check if there are correlations between variables before the factor analysis is conducted. Correlations between variables are important for forming common factors. However, too much correlation between all variables is not desirable, because then the variables may not uniquely contribute to the relevant factors. Table 19 in the appendix shows the correlations between variables used in factor analysis. We find that some of the variables are correlated more between each other than with others, which is desirable case. There are other tests, used to check whether data are suitable for factor analysis. Bartlett test of sphericity tells us if we have sufficient inter-correlations to conduct the analysis. The result of the test is significant with α 0.05, which indicates that there are sufficient inter-correlations (Table 20 in the appendix). Kaiser-Meyer-Olkin (KMO) measure shows us sampling adequacy. The statistic is a measure of the proportion of variance among variables that might be common variance, the lower the proportion, the more suited data is for factor analysis. KMO return values between 0 and 1. KMO values between 0.8 and 1 indicate the sampling is adequate. KMO values less than 0.6 indicate the sampling is not adequate and remedial action should be taken, values close to zero indicate that there are widespread correlations, which are a problem for factor analysis. Table 20 shows the results from these tests. KMO value of 0.801 indicates that our sample is adequate.

After concluding that our sample is suitable for factor analysis, we proceed with factor identification. We need to select which criterion will be used in factor extraction. One of the criterion used is called ‘Kaiser’s criterion’, which recommends retaining all factors that have the eigenvalue of 1 and above (Kaiser, 1960). Another criterion called ‘Jolliffe’s criterion’ suggests to retain factors above 0.70 (Jolliffe, 1986). It is important to select which criterion is most suitable depending on how many factors we expect. Humlum et al. (2012) identify two factors, which are determined by the number of eigenvalues greater than 1. As in Humlum et.al.(2012) we use Principal Factors method for factor extraction. However, we do not restrict ourselves to only 2 factors and therefore retain factors with the eigenvalues larger than 0.8, which determines 4 factors.(Table 9). In order to interpret factors, we need to use Kaiser’s Varimax procedure, which maximizes the variance of the squared loadings within each column and returns meaningful factor loadings.

The factor loadings support our expectations, 4 distinct factors emerge, which are named accordingly to which variables load mostly on them. As mentioned above, factor loadings greater than 0.3 are regarded as ‘significant ‘. As Lent et.al.(2001) suggest, factors influencing people decision making regarding college major might be both internal (self-efficacy and outcome expectation, coping efficacy) and external (social contextual barriers and supports).We identify a factor, which reflects the coping attitudes and we name it ‘Cope’. Note that coping factor loads on questions regarding statements such “If you are stressed out or find yourself in a difficult situation: you feel anxious about not being able to cope”. In this case the higher the coping score the higher tendency not to be able to cope with stressful situations. Next we identify a factor, which reflects self-esteem and name it ‘Esteem’. Self-esteem factor can be interpreted as self confidence in dealing with difficult problems and being able to solve them. Furthermore, as in Humlum et al. (2012) we identify ‘Social’ factor which can be interpreted as social orientation as it loads on statements about work values, such “It is important to pursue occupation in which I can help other people”. While ‘career’ factor can be interpreted as career orientation, which reflects on statements about work values such “To earn a lot of money, have a good wage”.

Table 1 Variable description

Variable name	Variable description	Factor loading
	“Social Factor”	
t2copa1	If you are stressed or find yourself in a difficult situation, you try to be with other people	0.32
t1vawi1	When you think of the future it is important to have a job, where you can always learn something new	0.49
t1vawi2	When you think of the future it is important to pursue an occupation in which you can fully deploy your competences	0.55
t1vawi3	When you think of the future it is important to have a job where you can be in touch with other people	0.60
t1vawi4	When you think of the future it is important to pursue occupation where you can help other people	0.59
t1vawi5	When you think of the future it is important to have a job which gives you a feeling of doing something sensible	0.51
	“Career Factor”	
t1vawe1	When you think of the future it is important to earn a lot of money	0.59
t1vawe2	When you think of the future it is important to have a secure position (security of unemployment)	0.47
t1vawe3	When you think of the future it is important to have a position with a lot of career opportunities	0.52
	“Esteem factor”	
t2seef1	I can always manage to solve difficult problems if I try hard enough	0.56

t2seef2	I am confident that I could deal efficiently with unexpected events	0.45
t2seef3	Thanks to my resourcefulness, I know how to handle unforeseen situations	0.53
t2seef4	I can usually handle whatever comes my way	0.52
t2copt2	I focus on a problem and see how I can solve it	0.35
t2copt4	I try to be organized so I can be on top of situation	0.40
t2pers2	If I decide to accomplish something I manage to see it through	0.62
t2pers3	I complete whatever I start	0.59
t2pers4	Even if I encounter difficulties I persistently continue	0.66
t2pers5	I even keep at a painstaking task until I have carried it through	0.65
	"Coping Factor"	
t2cope1	If you are stressed out or find yourself in a difficult situation: you get angry	0.52
t2cope2	If you are stressed out or find yourself in a difficult situation: you feel anxious about not being able to cope	0.63
t2cope3	If you are stressed out or find yourself in a difficult situation: you blame yourself for not knowing what to do	0.59
t2cope4	If you are stressed out or find yourself in a difficult situation: you wish you could change what happened	0.51

Each individual has a score on a factor. Factor scores are predicted for each individual and will be used as explanatory variables in the educational choice models. After naming and storing our factors we can check what scores individuals have on different factors. As in Humlum et al. (2012) we expect that individuals with high "Career factor" scores would tend to choose longer education paths, and with high "Social factor" would choose shorter. Also it is expected that females would have higher scores on social factor while male would have higher scores on career factor. The table below shows summary statistics for the factor scores. The 2002 sample results reveal that women on average have a smaller score on self-esteem factor than men, what is more, women score negatively on self-esteem factor while men score positively. Also women tend to blame themselves for not being able to cope with stressful situations more often than men, which is represented by the coping factor. Moreover, women have on average larger scores on social factor than on the career, while for men it is the opposite. Furthermore, females on average score negative on career factor and positive on social factor, while males on average score positive on career factor and negative on social factor. Thus, gender differences, which represent non-cognitive abilities are persistent in our sample of 2002.

Table 2 Factor Means

	Female		Male	
Variable	Mean	S.d.	Mean	S.d.
Career factor	-0.074	0.77	0.12	0.76
Social factor	0.12	0.77	-0.19	0.94
Esteem factor	-0.08	0.91	0.12	0.85
Coping factor	0.24	0.77	-0.38	0.76

5.2 Logit Analysis. Survey Wave 2002

Model 1:

In the first specification, we check what effect a gender dummy variable has on choosing certain path of education. The result shows that the marginal effect of a female dummy variable on choosing gymnasium is 0.21, which is statistically significant at 0.01 level.

Model 2:

In the second specification, cognitive abilities: math and reading scores are added. These scores are taken from the PISA 2000 base survey. The reading estimate is positively associated with choosing gymnasium, yielding a marginal effect of 0.002, which is significant at 0.01 level. The math score is also positively associated with choosing gymnasium and is significant at 0.05 level. After adding the cognitive measures the effect of female dummy variable remains very similar to that in the first specification.

Model 3:

In the third specification, social, career, self-esteem and coping ability factors are added. The result shows that social factor yields a marginal effect of 0.087 of choosing gymnasium, this effect is significant. While the marginal effect of career factor is negative, -0.12, and is significant. The reading estimate is positively associated with choosing gymnasium, which is significant. The math score is also positively associated with choosing gymnasium and is significant. Self-esteem and coping variables are positively associated with choosing gymnasium, however, their effects seem to be insignificant.

Model 4:

In the fourth specification, family background variables that are taken from PISA 2000 base survey are added. Mothers and fathers tertiary education dummy variable, increases marginal effect of probability of choosing gymnasium by 0.19 respectively, with both effects being significant. Family wealth seems to have positive marginal effect on choosing gymnasium, however, this effect is insignificant. The marginal effect of career factor in probability of choosing gymnasium remains unchanged, -0.12 and is significant. The effect of social factor remains unchanged. Coping factor becomes significant in probability of choosing gymnasium at 0.05 level. Given the Akaikes information criteria, the fourth model is preferred.

We test the predictability power of our last model. The result shows that the fourth model has correctly classified 74.03% of the cases. 229 out of all 337 cases were correctly classified as positive, while 328 out of total 414 cases were correctly classified as negative.

We also regard fourth specification separately for men and women. The results for females show that, mother's and father's tertiary education dummy variables increase the odds ratio of choosing gymnasium as compared to the base outcome-vocational education and are significant on a 0.05 level. The increase in the reading literacy value would also increase the odds ratio of choosing gymnasium as compared to vocational education path and this result is significant. The 'social' factor for females increase the odds ratio of choosing gymnasium, while 'career' factor decrease the odds, with both results being significant. When looking at the males, results are similar. Both mother's and father's tertiary education dummy variables increase the odds ratio for males choosing gymnasium as compared to vocational path, however, mother's educational dummy variable is not significant in male case. The 'social' factor for males increase the risk ratio of choosing gymnasium, while 'career' factor decrease the odds ratio, however only 'career' factor is significant for males when choosing gymnasium against the vocational path. Thus, we conclude that both 'social and 'career' factors are important for women when choosing gymnasium against vocational path, however for men only 'career' factor is important.

Table 3 Logit Analysis 2002 (Model 1,2,3,4)

Logit Analysis of the Choice of the Educational Path: Marginal Effects of Choosing Gymnasium

Variable	Gymnasium ME, z-statistic, s.e.	Gymnasium ME, z-statistic, s.e.	Gymnasium ME, z-statistic, s.e.	Gymnasium ME, z-statistic, s.e.
Female	0.21*** (7.03) (0.03)	0.20*** (5.61) (0.036)	0.15** (3.18) (0.047)	0.17*** (3.62) (0.048)
Math score		0.001** (3.29)	0.001** (3.15)	0.001** (2.94)

		(0.0003)	(0.0003)	(0.0004)
Reading score		0.002*** (7.46) (0.0003)	0.002*** (6.45) (0.0003)	0.002*** (6.47) (0.0003)
Esteem			0.012 (0.53) (0.023)	0.018 (0.73) (0.025)
Social			0.087*** (3.49) (0.025)	0.085** (3.17) (0.027)
Career			-0.12*** (-4.28) (0.028)	-0.12*** (-4.03) (0.03)
Cope			0.043 (1.63) (0.026)	0.059** (2.10) (0.028)
Mother tertiary education				0.19*** (3.54) (0.056)
Father tertiary education				0.19*** (4.16) (0.047)
Urban				0.136** (3.16) (0.043)
Family wealth				0.007 (0.25) (0.03)
Age				0.0001*** (5.29) (0.0002)
Pseudo R^2	0.034	0.168	0.194	0.249
Observations	1001	973	752	751

Note: Statistical significance at the 1 percent level***,5 percent level **, 10 percent level*

To summarize, we first regard a base model adding only female dummy variable. The result show that being a female is positively associated with choosing gymnasium and this result is statically significant. Next, we add cognitive abilities to our model. We find that both math and reading scores are positively associated with choosing gymnasium studies and are statically significant. Subsequently, we add social, career, esteem and coping factors. We are mainly interested in social and career factors. Career factor is negatively related to choosing gymnasium, while social factor is positively related to choosing gymnasium with both effects being significant. We gradually add more variables and the last model with background control and ability variables results in: The marginal effect of career factor in probability of choosing gymnasium becomes stronger and is negatively associated with choosing gymnasium. Social factor is positively associated. The reading and math scores are positively associated with choosing gymnasium. We check the results separately for men and women and it turns out that both 'social and 'career' factors are important for women when choosing gymnasium against vocational path, however for men only 'career' factor is important.

5.3 Constructing Factors. Survey Wave 2007

Before we proceed with the analysis of 2007 survey data, we repeat the same procedure with testing sampling adequacy as above in 2002 case. Table 17 shows the correlations between variables used for factor analysis. Bartlett test of sphericity tells us that we have sufficient inter-correlations to conduct the analysis since the test is significant with α 0.05. KMO value of 0.842 is even bigger than in the case of 2002 sample, which indicates that our sample is adequate. We retain 4 factors .The factor loadings are summarized below:

Table 4 Factor Loadings 2002

Variable name	Variable description	Factor loading
	"Social Factor"	
T7copa1	If you are stressed or find yourself in a difficult situation, you try to be with other people	0.58
T7copa2	If you are stressed or find yourself in a difficult situation, you buy something for yourself	0.34
T7copa4	If you are stressed or find yourself in a difficult situation, you visit a friend	0.62
T7vawi3	When you think of the future it is important to have a job where you can be in touch with other people	0.59
T7vawi4	When you think of the future it is important to pursue occupation where you can help other people	0.53
T7vawi5	When you think of the future it is important to have a job which gives you a feeling of doing something sensible	0.43
	"Career Factor"	

T7vawe1	When you think of the future it is important to earn a lot of money	0.64
T7vawe2	When you think of the future it is important to have a secure position (security of unemployment)	0.45
T7vawe3	When you think of the future it is important to have a position with a lot of career opportunities	0.63
T7vawe4	When you think of the future it is important to have a position which is recognized and respected by others	
T7vawi2	When you think of the future it is important to pursue an occupation in which you can fully deploy your competences	0.36
	"Esteem factor"	
T7seef1	I can always manage to solve difficult problems if I try hard enough	0.53
T7seef2	I am confident that I could deal efficiently with unexpected events	0.44
T7seef3	Thanks to my resourcefulness, I know how to handle unforeseen situations	0.44
T7seef4	I can usually handle whatever comes my way	0.42
T7copt1	I analyze the problem before reacting	0.32
T7copt2	I focus on a problem and see how I can solve it	0.34
T7copt4	I try to be organized so I can be on top of situation	0.37
T7pers2	If I decide to accomplish something I manage to see it through	0.62
T7pers3	I complete whatever I start	0.60
T7pers4	Even if I encounter difficulties I persistently continue	0.71
T7pers5	I even keep at a painstaking task until I have carried it through	0.73
	"Coping Factor"	
T7cope1	If you are stressed out or find yourself in a difficult situation: you get angry	0.48
T7cope2	If you are stressed out or find yourself in a difficult situation: you feel anxious about not being able to cope	0.69
T7cope3	If you are stressed out or find yourself in a difficult situation: you blame yourself for not knowing what to do	0.64

T7cope4	If you are stressed out or find yourself in a difficult situation: you wish you could change what happened	0.50
T7seef1	I can always manage to solve difficult problems if I try hard enough	-0.30
T7seef2	I am confident that I could deal efficiently with unexpected events	-0.39
T7seef3	Thanks to my resourcefulness, I know how to handle unforeseen situations	-0.45
T7seef4	I can usually handle whatever comes my way	-0.38

We check once again what scores individuals have on different factors. The pattern is consistent with the 2002 wave results. The results show that women have on average larger scores on social factor than on the career, while for men it is the opposite. Furthermore, females on average score negative on career factor and positive on social factor, while males on average score positive on career factor and negative on social factor. Men on average have larger self-esteem scores and are more confident than women in solving difficult problems. Coping factor for females on average is positive, meaning that females tend to stress more when facing difficult situations. This pattern is similar to the survey wave 2002 results and the gender difference in non-cognitive abilities is also persistent in our sample of 2007.

Table 5 Factor means 2007

	Female		Male	
Variable	Mean	S.d.	Mean	S.d.
Career factor	-0.01	0.80	0.01	0.82
Social factor	0.24	0.80	-0.4	0.79
Esteem factor	-0.01	0.89	0.01	0.89
Coping factor	0.2	0.84	-0.3	0.77

5.4 Logit and Multinomial Analysis. Survey Wave 2007

The 2007 wave may reveal a more complete picture of the education path chosen, than the 2002 wave, because in the 2007 wave respondents are at the age of 23 and are either enrolled in higher education or have completed education. However, 2007 survey questionnaire asks about current status of education. Thus, some individuals, who have finished education before 2007 would appear under category ‘no education’. Therefore, we aggregate the data from 2002 until 2007 in order to acquire the highest educational level achieved.

Model 1:

In the first specification, female dummy variable is negatively associated with choosing University, however, this effect is not statistically significant.

Model 2:

In the second specification cognitive abilities: math and reading scores are added. Higher reading and math scores are positively associated with choosing university studies. The marginal effect of reading score on choosing university studies is 0.003 and is significant at 0.01 level. The marginal effect of math score is 0.0015 and is also significant at 0.01 level.

Model 3:

In the third specification, social, career, coping and self-esteem factor variables are added. The size of marginal effect of social factor is 0.052, however, it is insignificant. The marginal effect of career factor is -0.04, and it is insignificant. These results are not in line with our expectations. As in Humlum et.al. (2012) we expect that the higher scores on career factor would induce individuals choosing longer education paths, and thus choose university studies, while higher scores on social factor would be negatively associated with choosing university studies. The marginal effect of self-esteem on choosing university is positive. This effect is in line with our expectations because, one would expect the individual, who is persistent and have high self-esteem, manages to solve difficult problems and complete tasks and thus would more probably choose higher education level, however, self-esteem factor is not statistically significant. Coping is positively associated with choosing university studies and is statistically significant on 0.05 level. Math and reading scores remain positively associated with choosing University studies and are statistically significant on 0.01 level.

Model 4:

In the fourth specification, family background variables are added. Mothers tertiary education dummy variable, increases marginal effect of probability of choosing university by 0.19, which is significant. Fathers tertiary education dummy variable, increases marginal effect of probability of choosing university by 0.22, which is also significant. Family wealth have positive marginal effect on probability of choosing university, and it is significant. Career factor is negatively associated with choosing university, while coping, self-esteem and social factors are positively associated with choosing university, although, their marginal effects are statistically insignificant. Given the Akaikes information criteria, the fourth model is preferred.

We test the predictability power of our last model. The result shows that the fourth model has correctly classified 74.04% of the cases. 191 out of all 267 cases were correctly classified as positive, while 233 out of total 298 cases were correctly classified as negative.

Next, as in 2002 survey wave, we regard fourth specification separately for men and women. The results for females show that, mother's and father's tertiary education dummy variables increase the odds of choosing university as compared to the base outcome-vocational education and are significant. The increase in the reading literacy value would also increase the odds of choosing university as compared to vocational education path and this result is significant. The 'social' factor for females increase the odds of choosing university, while 'career' factor decrease the odds, however, these results are not statistically significant. When looking at the males, results are similar. Both mother's and father's tertiary education dummy variables increase the odds for males choosing university as compared to vocational path, however, mother's educational dummy variable is not significant in male case. The increase in the reading literacy value would also increase the odds of choosing university as compared to vocational education path and this result is significant. The 'social' factor for males increase the odds of choosing university, while 'career' factor decrease the odds, however both factors are not statistically significant. Thus, we conclude that both 'social and 'career' factors are not that important for women and men when choosing university against vocational path.

Table 6 Logit Analysis 2007 (Model 1,2,3,4)

Logit Analysis of the Choice of the Educational Path: Marginal Effects of Choosing University

Variable	University M.E., z-statistic, s.e.	University M.E., z-statistic, s.e.	University M.E., z-statistic, s.e.	University M.E., z-statistic, s.e.
Female	-0.005	-0.005	-0.07	-0.059

	(-0.30) (0.019)	(-0.09) (0.055)	(-1.14) (0.06)	(-0.91) (0.066)
Reading Score		0.003*** (6.43) (0.0004)	0.003*** (6.47) (0.0004)	0.003*** (6.55) (0.0004)
Math Score		0.0015*** (3.65) (0.0004)	0.002*** (3.71) (0.0004)	0.002*** (3.61) (0.005)
Esteem			0.01 (0.33) (0.03)	0.0002 (0.01) (0.033)
Social			0.052 (1.57) (0.033)	0.046 (1.32) (0.035)
Career			-0.04 (-1.29) (0.031)	-0.043 (-1.26) (0.034)
Cope			0.061** (1.98) (0.031)	0.084** (2.53) (0.033)
Mother Tertiary education				0.19** (2.73) (0.068)
Father Tertiary education				0.22*** (3.86) (0.056)
Family wealth				0.12** (3.00) (0.041)
Urban				0.034 (0.64) (0.05)
Age				0.0002*** (5.05) (0.0004)
Pseudo R^2	0.000	0.19	0.203	0.277
Observations	2902	570	566	565

Note: Statistical significance at the 1 percent level***,5 percent level **, 10 percent level*

To sum up: In the first specification, we consider a base model, where we only add the female dummy variable. We find that the female dummy variable is negatively associated with choosing university studies. Next, we add cognitive abilities to our model. We find that both math and reading scores are positively associated with choosing university studies and are statically significant. Following we add career, social, coping and self-esteem factors to the model. Social factor is positively associated with choosing university studies, while career factor is negatively associated with choosing university studies. These results are not in line with our expectations. As in Humlum et.al. (2012) we expect that the higher scores on career factor would induce individuals choosing longer education paths, and thus choose university studies, while higher scores on social factor would be negatively associated with choosing university studies. We gradually add more variables and in the final model, we have background control variables and cognitive ability variables. The results from the final model show that career factor is negatively associated with choosing university, while coping, self-esteem and social factors are positively associated with choosing university, although, their marginal effects are statistically insignificant. We check the results separately for men and women and it turns out that the social factor for males increase the odds of choosing university, while career factor decrease the odds, however both factors are not statistically significant. Thus, we conclude that both ‘social’ and ‘career’ factors are not that important for women and men when choosing university against vocational path.

In TREE survey, participants are asked ‘What kind of job do you have presently? What kind of work do you perform at this job?’. The answers to this question are coded according to the International Standard Classification of Occupations revised in 1987 (ISCO 88). An occupational classification is a tool for organizing all jobs in an establishment, an industry or a country into a clearly defined set of groups. Occupational classifications are usually designed to serve several purposes. Legislators use occupational statistics in formation of policies, while different researchers use this information in their analysis of social differences, behaviours, earnings, etc. In the context of ISCO a job is defined as a set of tasks and duties which are carried out by one person. In ISCO-88 occupation are grouped together and further aggregated mainly on the basis of the similarity of skills required to fulfil the tasks and duties of the jobs. ISCO-88 defines four levels of aggregation, consisting of (source OIT):

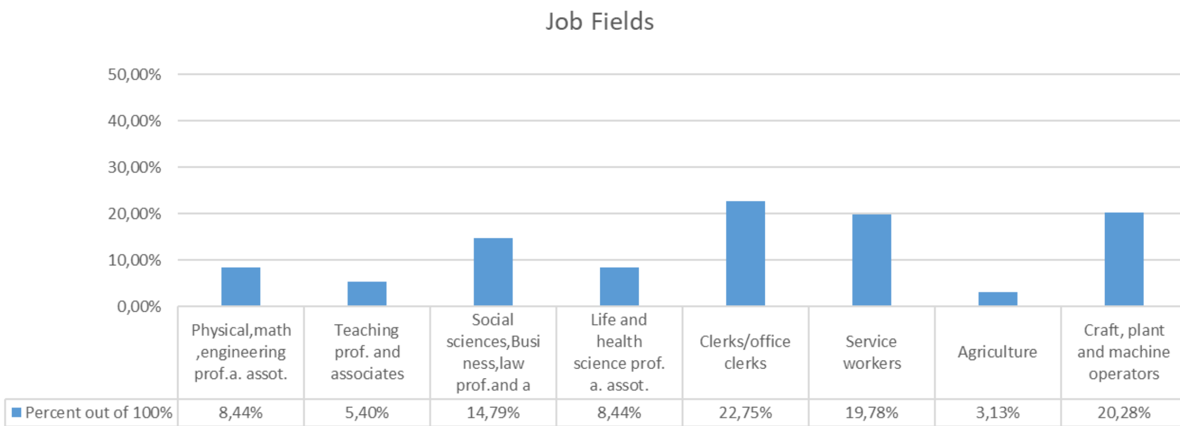
- 10 major groups
- 28 sub-major groups (subdivisions of major groups)
- 116 minor groups (subdivisions of sub-major groups)
- 390 unit groups (subdivisions of minor groups)

Next we will analyse the choice of field of occupation. As described above, the occupation fields are grouped into 10 major groups. For the sake of our analysis the occupations will be further aggregated into 8 categories as follows: 'Physical, Math, Engineering professionals and associates', 'Teaching professionals and associates', Social sciences, Business, Law professionals and associates', 'Office clerks', 'Service workers', 'Agriculture', 'Craft, Plant and Machine operators'. Major group in ISCO classification are also delineated with the reference to the four broad skills levels. After aggregating occupations into 8 categories, the skills reference become:

Table 7 Occupation groups

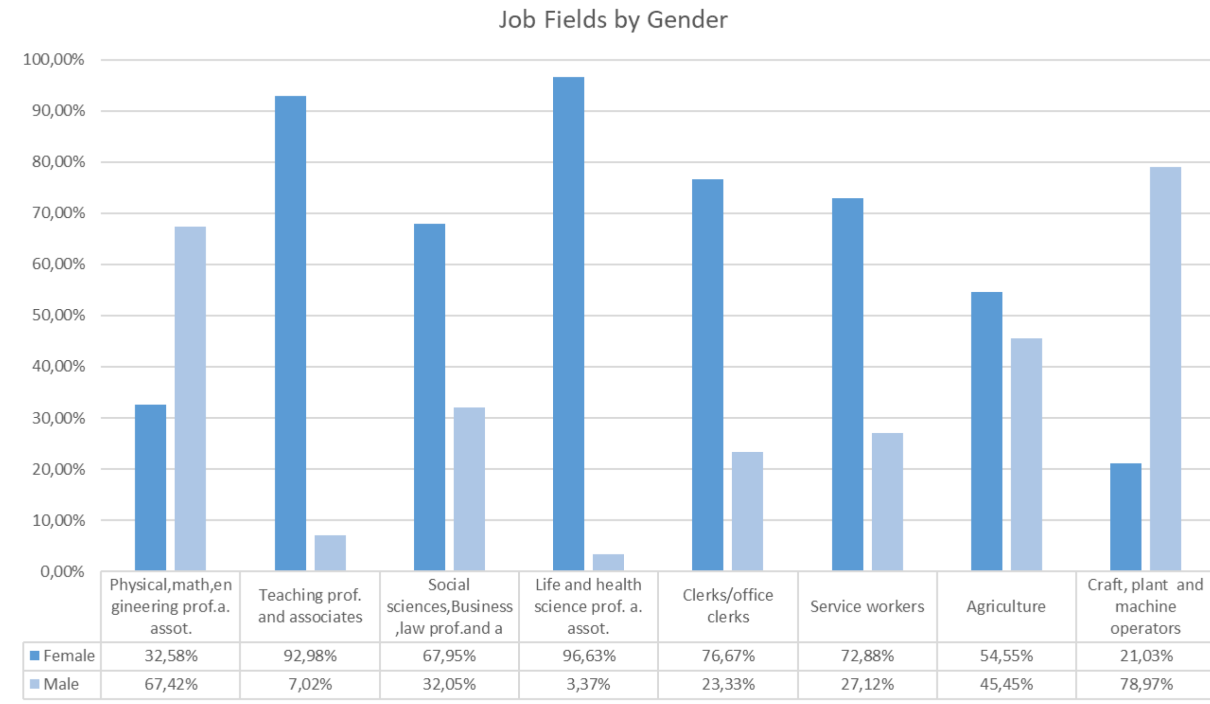
Occupation group	**Skill level**
Physical, Math, Engineering professionals and associates	4th and 3rd
Teaching professionals and associates	4th and 3rd
Social sciences, Business, Law professionals and associates	4th and 3rd
Life and Health science professionals and associates	4th and 3rd
Clerks/Office clerks	2nd
Service workers	2nd
Agriculture	2nd
Craft, Plant and Machine operators	2nd

The summary in a graph below shows that majority (23%) of the individuals work as Clerks/office clerks. To this category belong: Secretaries, numerical clerks such as accounting and bookkeeping clerks, material-recording and transport clerks, library, mail related clerks and etc. The second largest group of individuals (20%) are employed as 'Craft, Plant and Machine operators'. This group include: Extraction and building trades workers, metal, machinery and related workers, blacksmiths, handicraft workers and etc. 17% of individuals are service workers, which are: travel attendants, housekeeping, restaurant services workers, personal care workers, protective service workers such as police officers etc. 15% of individuals are employed as Social sciences, Business, Law professionals and associates, such as economists, lawyers, psychologists, authors etc.8% are employed as 'Physical, Math, Engineering professionals and associates', 8% as 'Life and Health science professionals and associates', 5% as 'Teaching professionals and associates'.



Picture 3 Percentage of individuals in each job field

If we look at the males and females picture separately, the results reveal that most of the women (28%) are employed as 'Clerks and office clerks', while most of the men (42%) as 'Craft, Plant and Machine operators' respectively. 16% out of all females, and 12% out of all males are employed as 'Social sciences, Business, Law professionals and associates'. 15% out of males are employed as Physical, Math, Engineering professionals and associates', while only 4% out of all women are employed in this field. While looking at the total population, 3% of females and 6% of males are employed in the Physical, Math and Engineering field. 8% out of all females are employed as 'Teaching professionals and associates', while only 1% out of all males are employed in this field. Looking at the total population 5% of women are employed in the Teaching field, whereas, only 0.4% of men work in teaching field. 13% percent out of women are employed in 'Life and Health science professionals and associates', while only 1% out of all men are working in this field. The picture of Swiss sample is similar to the overall European picture as statistics indicate. Majority of females work in the service field, as office clerks and in the social sciences, business and law occupations. Whereas, majority of men are working as craft, plant and machine operators, Physical, Math and Engineering occupations and also in the Social sciences, Business and Law occupations. What is more, the result show that there are half as many males as females in the Physical, Math and Engineering occupations. Thus, there is gender disparity within certain professions.



Picture 4 Gender division by occupation fields

Next, we analyse the results from the model of the choice of occupation field applying multinomial logit analysis:

Model 5:

In the 5th specification, female dummy variable is added in the occupation field model. The result shows that being a female is negatively associated with choosing physics, math and engineering professions and this result is statistically significant. However, being a female is positively associated with choosing, teaching professions, life and health science professions, and social sciences, business and law occupations.

Model 6:

In the 6th specification, we add ability measures. Note that the limitation of this specification is that when adding the ability variables the number of observations that we have decreases considerably. The marginal effect of reading score on probability of choosing Physics, math and engineering professions is positively related, however, is very small 0.0006 and insignificant. The marginal effect of math score on probability choosing this field is also positively related but very small 0.0005 and is significant at 0.1 level. Being a female reduces marginal effect on probability of choosing this field by -0.084, this effect is significant. The marginal effects on probability of choosing Teaching professions are as follows. The reading score is positively associated with choosing these professions. The marginal effect of reading score is 0.0008 and it is significant. The math score is negatively associated with choosing these professions. The marginal effect of math score is -0.0003, which is insignificant. Being a female increases marginal probability of choosing this field by 0.04. The result for 'Social sciences, Business and Law' professions show that reading score is positively associated with choosing these occupations. The marginal effect of reading score is 0.0009, which is significant on the alpha level of 0.1. The marginal effect of math score is negative, -0.0001, and is insignificant.

Model 7:

In the 7th specification, we add social, career, self-esteem and coping factors. Adding the identity factors, decrease the marginal effect of female dummy variable to -0.035 on choosing Physics, math and engineering professions. However, female dummy variable remains statistically insignificant. The effect of cognitive variables remain almost unchanged, with only math variable being statistically significant on choosing Physics, math and engineering professions. The effect of career, social and coping factors on choosing Physics, math and engineering professions is negative, however only career and coping factor is statistically significant. In case of choosing teaching professions, adding the identity factors, slightly decrease the effect of female dummy variable. However, the marginal effect of reading score slightly increases. In case of choosing the teaching profession, social factor is positively associated, as it is expected. The marginal effect of social factor is 0.026 and it is statistically significant. The marginal effect of career factor is negative, however, its effect is insignificant. In case of life and health science professions, the female dummy variable is positively associated with choosing these professions and is statistically significant. Self-esteem, career and social factors are also positively associated with choosing these professions, however, only social factor is statistically significant. In case of social sciences, business and law professions, female dummy variable is positively associated with choosing these professions. However, elf-esteem, career, social, and coping factors are all negatively associated with choosing these professions.

Model 8:

In the 8th specification, we add parent tertiary education, wealth and agglomeration variables. The results from choosing physics, math, engineering professions show that: Mother and father tertiary education dummy is negatively associated with choosing these professions, with marginal effects being -0.02 and -0.07 respectively. Family wealth is positively associated with choosing these professions. Social and career factors remain negatively associated with choosing these professions. Being female is negatively associated with choosing these professions. Only female dummy variable, career factor and father tertiary education are statistically significant in case of choosing physics, math and engineering professions. The result for 'social sciences, business and law' professions show that mother tertiary education is positively associated with choosing these professions, while father tertiary education is negatively associated. The math score, career, social, esteem and coping factors are all negatively associated with choosing these professions. However, these effects are insignificant. Reading score and being a female are positively associated and these variables are significant. In case of life science and health sciences professions, being a female is positively associated with choosing these professions and this results is statistically significant. Social factor is positively associated while career factor is negatively associated with choosing these professions, however, only social factor is statistically significant. The results for teaching professions show that being a female is positively associated with choosing these professions. Social factor is positively associated, while career factor is negatively associated with choosing these professions. However, these effects are insignificant. The results for the highest skill level occupations are summarized below. The rest of the results can be found in the appendix. (Tables 123-130)

Table 8 Multinomial Logit Analysis 2007 (5)

Table: Multinomial Logit Analysis of the Choice of Job Field

5) Variable	Physics, math, engineering professions M.E., (z-statistic), (s.e.)	Teaching professions M.E., (z-statistic), (s.e.)	Life and health science professions M.E., (z-statistic), (s.e.)	Social sciences, business and law M.E., (z-statistic), (s.e.)
Female	-0.10*** (-5.33) (0.019)	0.07*** (6.07) (0.012)	0.13*** (8.95) (0.014)	0.04* (1.81) (0.023)

Pseudo R^2 = 0.076
Observations 1055

Note: Statistical significance at the 1 percent level***,5 percent level **, 10 percent level*

Table 9 Multinomial Logit Analysis 2007 (6)

Table: Multinomial Logit Analysis of the Choice of Job Field

6) Variable	Physics, math, engineering professions M.E., (z-statistic), (s.e.)	Teaching professions M.E., (z-statistic), (s.e.)	Life and health science professions M.E., (z-statistic), (s.e.)	Social sciences, business and law M.E., (z-statistic), (s.e.)
Female	-0.084** (-2.04) (0.041)	0.04 (1.49) (0.027)	0.98** (3.37) (0.029)	0.014 (0.27) (0.051)
Math score	0.0005* (1.89) (0.0003)	-0.0003 (-1.16) (0.0003)	0.0000 (0.28) (0.0000)	-0.0001 (-0.33) (0.0005)
Reading score	0.0006 (2.16) (0.0003)	0.0008** (2.97) (0.0003)	0.0000 (1.49) (0.0000)	0.0009* (2.06) (0.0005)

Pseudo $R^2 = 0.18$
Observations 216

Note: Statistical significance at the 1 percent level***,5 percent level **, 10 percent level*

Table 10 Multinomial Logit Analysis 2007 (7)

Table: Multinomial Logit Analysis of the Choice of Job Field

7) Variable	Physics, math, engineering professions M.E., (z-statistic), (s.e.)	Teaching professions M.E., (z-statistic), (s.e.)	Life and health science professions M.E., (z-statistic), (s.e.)	Social sciences, business and law M.E., (z-statistic), (s.e.)
Female	-0.035 (-0.86) (0.04)	0.01 (0.71) (0.014)	0.06** (2.00) (0.032)	0.05 (0.64) (0.07)
Esteem	0.004 (0.19) (0.02)	-0.014 (-1.53) (0.01)	0.00001 (0.18) (0.00007)	-0.015 (-0.37) (0.04)
Social	-0.03 (-1.02) (0.029)	0.026* (1.74) (0.015)	0.0003** (3.08) (0.0001)	-0.022 (-0.45) (0.05)
Career	-0.05* (-1.83) (0.026)	-0.02 (-1.43) (0.012)	0.00001 (0.11) (0.0001)	-0.05 (-0.84) (0.06)
Cope	-0.05* (-1.65) (0.028)	0.005 (0.52) (0.01)	-0.00001 (-0.21) (0.00006)	-0.04 (-0.98) (0.04)
Reading Score	0.0005 (1.64) (0.0003)	0.001* (1.69) (0.0002)	0.0000 (1.17) (0.0000)	0.001** (2.00) (0.0004)
Math Score	0.0005* (1.8) (0.0003)	-0.0001 (-0.66) (0.0002)	0.0000 (0.4) (0.0000)	-0.0001 (-0.26) (0.0005)

Pseudo $R^2 = 0.239$
Observations 215

Note: Statistical significance at the 1 percent level***,5 percent level **, 10 percent level*

Table 11 Multinomial Logit Analysis 2007 (8)

Multinomial Logit Analysis of the Choice of Job Field

8) Variable	Physics, math, engineering professions M.E., (z-statistic), (s.e.).	Teaching professions M.E., (z-statistic), (s.e.)	Life and health science professions M.E., (z-statistic), (s.e.)	Social sciences, business and law M.E., (z-statistic), (s.e.)
Female	-0.034 (-0.91) (0.037)	0.012 (0.76) (0.016)	0.053* (1.86) (0.029)	0.048 (0.65) (0.074)
Esteem	0.008 (0.39) (0.19)	-0.01 (-1.25) (0.01)	0.00002 (0.47) (0.00005)	-0.017 (-0.38) (0.046)
Social	-0.033 (-1.16) (0.024)	0.018 (1.31) (0.014)	0.0002** (2.04) (0.00008)	-0.02 (-0.39) (0.05)
Career	-0.053** (-2.21) (0.024)	-0.013 (-0.99) (0.013)	0.0000 (0.05) (0.00008)	-0.058 (-0.92) (0.063)
Cope	-0.05 (-1.52) (0.033)	0.003 (0.35) (0.008)	0.0000 (0.20) (0.00004)	-0.05 (-1.18) (0.04)
Reading Score	0.0006 (1.53) (0.0004)	0.0003 (1.37) (0.0002)	0.0000 (1.22) (0.0000)	0.001** (2.10) 0.0005
Math Score	0.0004 (1.37) (0.0003)	-0.00013 (-0.68) (0.0002)	0.0000 (0.13) (0.0000)	-0.0001 (-0.12) (0.0005)
Mother tertiary education	-0.024 (-0.35) (0.067)	0.032 (0.84) (0.039)	0.0000 (0.04) (0.0002)	0.033 (0.32) (0.1)
Father tertiary education	-0.07** (-1.99) (0.036)	0.058 (1.61) (0.036)	-0.00001 (-0.14) 0.0001	-0.06 (-0.85) (0.075)
Urban	-0.023 (-0.65)	-0.007 (-0.37)	-0.0002* (-1.82)	0.07 (1.08)

	(0.036)	(0.018)	(0.0001)	(0.07)
Wealth	0.022	0.013	0.00005	0.003
	(0.78)	(-1.17)	(0.13)	(0.07)
	(0.028)	(0.011)	(0.00008)	(0.04)

Pseudo R^2 = 0.295
Observations 215

Note: Statistical significance at the 1 percent level***,5 percent level **, 10 percent level*
n.a. very negligible effect

To summarize: In this section we looked at the chosen occupation fields. In the first specification we female dummy variable. It is found that being a female is negatively associated with choosing physics, math and engineering professions and this result is statistically significant. However, being a female is positively associated with choosing, teaching professions, life and health science professions, and social sciences, business and law occupations. In the second specification reading and math scores were added. The result showed that math is positively associated with choosing Physics, math and engineering professions, while reading score is positively associated with teaching professions, life and health science professions, and social sciences, business and law occupations as expected. In the third specification identity factors were added. The effect of career, social and coping factors on choosing physics, math and engineering professions is negative, however only career and coping factor is statistically significant. In case of choosing teaching professions, adding the identity factors, slightly decrease the effect of female dummy variable. In case of choosing the teaching profession, social factor is positively associated, as it is expected. The marginal effect of career factor is negative, however, its effect is insignificant. In case of life and health science professions, the female dummy variable is positively associated with choosing these professions and is statistically significant. Self-esteem, career and social factors are also positively associated with choosing these professions, however, only social factor is statistically significant. In case of social sciences, business and law professions, female dummy variable is positively associated with choosing these professions. However, elf-esteem, career, social, and coping factors are all negatively associated with choosing these professions. We gradually added more explanatory variables, however, the number of individuals we have decreased considerably. In the last specification, we added background control variables. The results from choosing physics, math, and engineering professions show that social and career factors remain negatively associated with choosing these professions. Being female is negatively associated with choosing these professions. Only female dummy variable, career factor and father tertiary education are statistically significant in case of choosing physics, math and engineering professions. The result for ‘Social sciences, Business and Law’ professions show that both career and social factor become negatively associated with choosing these professions. The reading score is positively associated, is significant in this case, while math score is negatively associated, and is insignificant in choosing these professions. In case of choosing teaching professions, career factor is negatively, while social factor is positively associated with these professions. However both effects are insignificant.

6. Conclusion

In order to address the gender gap in certain occupations it is important to recognize the social influences as well as economical reasoning in the education and later career choices. In this paper we consider a model based on Humlum et al. (2012), where choice of education path and occupation field is motivated by pay off in terms of identity due to more rewarding self-image.

This thesis use Swiss data set TREE and PISA 2000 surveys, which together provide information on individuals and their parents and also contain information on individual attitudes towards work. Using factor analysis, which is based on a range of questions about self-perception and attitudes, we construct four factors which are interpreted as social orientation, career orientation, self-esteem and coping ability. Women on average have higher score on social factor while men have a higher score on career factor.

Consequently, we test empirically whether these factors matter for education path and occupation field choice. We use multinomial logit analysis of choice of education path/career field using the four factors and other control variables. We conduct our analysis based on 2002 and 2007 survey waves.

Based on 2002 wave results we find that social, career and coping factors are important for educational path choice. When looking separately at men and women we find that both social and career factors are important for women when choosing gymnasium against vocational path, while for males only career factor matters.

The results from 2007 wave show that with respect to choosing university against vocational path, career factor is negatively associated with choosing university, while coping, self-esteem and social factors are positively associated with choosing university, although, their marginal effects are statistically insignificant. When looking separately at men and women, we find that social factor increase the odds of choosing university, while career factor decrease the odds for both genders, however, these results are not statistically significant.

In case of choosing occupation field, we find that individuals with higher social factor tend to choose teaching and life and health science professions, and these results are statistically significant. While career factor is negatively associated with choosing engineering, teaching and social sciences professions, but its effects seems to not matter except in case of physics, math, engineering professions. We also find that people with higher self-esteem tend to choose engineering, teaching and social sciences professions, however this results is insignificant.

It appears that these results are somewhat undetermined, however, this is what we know with this degree of uncertainty based on this study. In some cases personal factors matter more in choosing educational paths/occupations, in other cases they do not matter. Nevertheless, the effects of self - perception and self-image, besides the cognitive abilities, should be taken into consideration, because the results show that there are gender differences in non-cognitive abilities.

The results from this study suggest that there are gender differences in self-perception and work attitudes. Cognitive skills play important role in the education major selection. Even though, female students often outperform male students in schools in terms of cognitive abilities, fewer women are choosing STEM majors that require high cognitive skills. As Lent and Brown (2001) suggest, there are external (social contextual) and internal (self-efficacy and outcome expectation, coping efficacy) barriers and supports in choosing college majors. Student may believe that s/he has strong math capabilities (task self-efficacy) yet lack confidence at withstanding negative peer pressure linked to pursuing math related major. This would imply that different information policies for attracting individuals to professions, where the gender gap is considerable high, such as STEM occupations in case of low women representation or teaching or health care in case of low men representation in those professions are needed. Interventions should focus not only on pecuniary benefits but also on stereotypes and aim at confidence and ability perceptions.

We recommend future research to focus on social influences in addition to economical reasoning in the education and later career choices on a wider scale. Getting a more balanced share of women and men in certain occupations and higher positions is important in closing gender power gap. Our results from Swiss sample shed some light on how self-perception and attitudes influence education paths/career choices. However, education system in Switzerland, with its strong focus on vocational training, differs from many of the Europe’s education systems. Our research lacks on a larger sample with more observations. Therefore, there is a need for further research.

7. References

Aigner, D. J., & Cain,G. G. (1977). Statistical theories of discrimination in labour markets. *Industrial & labour Relations Review*, *Vol.30No.2*, p.175-187.

Akerlof, G. & Kranton, R. (2000). Economics and identity. *Quarterly Journal of Economics, Vol. 115, No. 3,* p. 715–753.

Akerlof, G. & Kranton, R. (2002). Identity and schooling: Some lessons for the economics of education. *Journal of Economic Literature, Vol. 40, No.*4, p. 1167-1201.

Bartholomew, D., Knott, M., Moustaki, I., (2011). Latent variable models and factor analysis: A unified approach. (3rded.). UK: John Wiley & Sons, Ltd.

Barrick, M. R., & Mount, M. K. (1991). The big five personality dimensions and job performance: A meta-analysis. *Personnel Psychology, Vol.44, No.1*, p.1-26.

Boskin, M.J. (1974). A conditional logit model of occupational choice. *Journal of Political Economy*, *Vol.82, No. 2*, p.389-398

Becker G. S. (1957). The Economics of Discrimination. (2nd ed.). Chicago: University of Chicago Press.

Burke, P. J., & Reitzes, D. C. (1981). The link between identity and role performance. *Social Psychology Quarterly*, *Vol.44*, *No. 2,* p.83-92.

Directorate-General for Justice and Consumers (European Commission). (2018). Report on equality between women and men in the EU 2018. Retrieved from: https://op.europa.eu/s/nxHZ

Cattel, R. B. (1945). The description of personality: Principles and findings in a factor analysis. *The American Journal of Psychology, Vol. 58, No.1,* p. 69-90.

Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology, Vol.54*, p.1–22.

Child D. (2006). The essentials of factor analysis. (3rd ed.). London, New York: Continuum International Publishing Group.

Fabrigar, L. R. & Wegener, D. T. (2011). Exploratory factor analysis. New York: Oxford University Press.

European Commission. (2019). Module 1: The race and framework directives. Retrieved from: https://www.era-comm.eu/anti-discri/e_learning/module1_1.html

Eurostat. (2018). Jobs still split along gender lines. Retrieved from: https://ec.europa.eu/eurostat/web/products-eurostat-news/-/EDN-20180307-1

Eurostat. (2019). Women in science and technology. Retrieved from: https://ec.europa.eu/eurostat/web/products-eurostat-news/-/EDN-20190211-1

European Parliament. (2018). More women in ICT: empowering women in the digital world. Retrieved from: https://www.europarl.europa.eu/news/en/headlines/society/20180301STO98927/more-women-in-ict-empowering-women-in-the-digital-world

Fietze, S., Holst, E., Tobsch, V. (2011). Germany's next top manager: Does personality explain the gender career gap? *Management Revue, Vol. 2, No.3,* p. 240-273.

Greene, W. H. (1951). Econometric analysis. (7thed.). New York: Prentice Hall.

Harding, S. & Hintikka, M. B. (1983). Discovering reality: Feminist perspectives on epistemology, metaphysics, methodology, and philosophy of science. Boston: D. Reidel Publishing.

Harman, H. H. (1976). Modern factor analysis. (3rded.). Chicago: University of Chicago Press.

Heckman, J. J., Stixrud, J., Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labour market outcomes and social behaviour. *Journal of labour Economics*. *Vol.24*, No.3, p.411-482.

Hoffmann, E. & Scott, M. (1993). The revised international standard classification of occupation. Retrieved from: http://www.oit.org/public/english/bureau/stat/download/papers/short.pdf

Humlum, M. K., Kleinjans, K. J., Nielsen, H. S. (2012). An economic analysis of identity and career choice. *Economic Inquiry, Vol.50, No.1,* p.39-61.

Jolliffe, I. T. (1972). Discarding variables in a principal component analysis, I: Artificial data. *Applied Statistics, Vol.21*, p.160-173.

Kaiser, H.F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, *Vol. 20*, p.141-151.

Lent, R., & Brown, S., (2001). The role of supports and barriers in the choice of Math/Science options: A test of social cognitive hypotheses. *Journal of Counseling Psychology*, *Vol.48, No.4*, p.474-483.

McDougall W. (1932). Of the words character and personality. *Journal of Personality*, *Vol. 1*, p. 3-16.

McFadden, D. (1973) Conditional Logit Analysis of Qualitative Choice Behaviour. In: Zarembka, P., Ed., p. 105-142, *Frontiers in Econometrics*, Academic Press.

Mercer (2016) Global report: When women thrive. Retrieved from: https://www.se.mercer.com/vara-publikationer/when-women-thrive-2016-report.html

Phelps, E. S. (1972). The statistical theory of racism and sexism. *The American Economic Review, Vol. 62, No. 4,* p. 659–661.

Spearman, C. (1904). General intelligence. Objectively determined and measured. *The American Journal of Psychology, Vol.15, No.2,* p.201-292.

Siann, G., & Callaghan, M. (2001). Choices and barriers: Factors influencing women's choice of higher education in science, engineering and technology. *Journal of Further and Higher Education*, *Vol.25, No.1*, p.85-95.

Swiss Education. (2019). Compulsory education. Retrieved from: https://swisseducation.educa.ch/en/compulsory-education-0

Swiss Education. (2019). Primary level. Retrieved from: https://swisseducation.educa.ch/en/primary-education

Swiss Education. (2019). Lower-secondary level. Retrieved from: https://swisseducation.educa.ch/en/lower-secondary-education

Swiss education (2019). Upper-secondary level. Retrieved from: https://swisseducation.educa.ch/en/upper-secondary-level-0

Swiss Education. (2019). The Swiss Education System Retrieved from: https://swisseducation.educa.ch/en/swiss-education-system-3

Swiss Education. (2016). Digital image. Retrieved from: https://swisseducation.educa.ch/en

Transitions from Education to Employment (TREE), Cohort 1 – 2000-2014 [Dataset]. University of Bern. Distributed by FORS, Lausanne, 2016, version 1.0.0. Retrieved from: https://doi.org/10.23662/FORS-DS-816-6

Willis, R. J., & Rosen, S. (1979).Education and self-selection. *Journal of Political Economy*, *Vol.87,* p.S7-S36.

Wolbers, M. H. J. (2000). The effects of level of education on mobility between employment and unemployment in the Netherlands. *European Sociological Review*, *Vol.16, No. 2,* p.185-200.

Yong,A. G., & Pearce, S. (2013). A beginner's guide to factor analysis: Focusing on exploratory factor analysis. *Tutorials in Quantitative Methods for Psychology*. Vol. 9, No. 2, p. 79-94.

8. Appendix

8.1 Education System in Switzerland

Every child in Switzerland has to complete the compulsory education. The total period of compulsory education amounts to eleven years. Primary level – including two years of kindergarten or a first learning cycle – comprises eight years. Lower secondary level takes three years. Children in these grades tend to be between 12 and 15. There is no nationwide exam at the end of ninth grade – the final year – so students receive no graduation certificate.

After completing the compulsory lower secondary education students may transfer to the upper secondary education. More than 90% of young people complete an upper secondary level programme. The adolescents complete upper secondary level at the age of 18/19 and receive a corresponding certificate.

Upper secondary education can be subdivided into general education programmes (1), and vocational education and training (VET) programmes (2).

1) The general education programmes include the Baccalaureate schools and the upper secondary specialised schools. They do not lead to professional qualifications, but prepare for tertiary level education programmes.

At the end of their senior high school studies, students must do a type of thesis as well as pass a series of examinations that, if successfully completed, result in a matura - high school leaving certificate - that allows admission to cantonal universities and Federal Institutes of Technology

2) Vocational education and training (VET), in which adolescents learn a profession, is mostly completed at training companies (apprenticeship) combined with teaching at a VET school. It can also be completed at full-time vocational schools.

The majority of adolescents enrol in vocational education and training (VET) after lower secondary level. There are VET programmes for some 250 different professions. In Switzerland, many professional qualifications are obtained in upper secondary level, while in other countries the same qualifications are obtained in tertiary level education. The Swiss system therefore differs from most foreign systems of vocational and professional education and training. VET is predominantly based on a dual system: practical training (apprenticeship) on three to four days at a training company is supplemented by theoretical classes (vocational and general educational subjects) on one to two days at the VET school.

Vocational education and training (VET) offers the following programmes:

- Two-year vocational education and training VET programme with Federal VET Certificate. The two-year vocational education and training VET programme leading to a

Federal VET Certificate offers adolescents with lower learning performance a federally recognised professional qualification. It prepares them for a less-demanding occupation.
- Three- or four-year VET programme with Federal VET Diploma. The three- or four-year VET programme leading to a Federal VET Diploma provides training for work in a particular profession.

Tertiary level education can be completed at universities or at professional education institutions. In Switzerland there are following universities:

- Universities: cantonal universities and Federal Institutes of Technology (ETH)
- Universities of applied sciences
- Universities of teacher education

Tertiary level professional education enables (in particular) professionals who have completed vocational education and training (VET) to specialise and to enhance their skills and knowledge. Tertiary level professional education offers the following educational programmes:

- Federal Diploma of Higher Education and Advanced Federal Diploma of Higher Education Examinations
- Colleges of Higher Education diplomas

8.2 Additional Tables

Table 1

List of Variables

Variable name	Description
Edul (2002)	Categorical variable of educational level of sample 2002 :1-Vocational, 2-gymnasium;3- no education
Edul (2007)	Categorical variable of educational level of sample 2007 :1-Vocational, 2-university; 3- no education
jobfield	Categorical variable of occupation field of sample 2007 :1- Physical, Math, Engineering professionals and associates, 2- Teaching professionals and associates; 3- Social sciences, Business, Law professionals and associates;4-Life and Health science professionals and associates;5- Clerks/Office clerks;6-Service workers;7- Agriculture; 8- Craft, Plant and Machine operators
Wealth	Index of family wealth from PISA 200 study
Urban	Agglomeration dummy 1-urban,0-rural
Wleread	Warm estimate of reading from PISA 200 study
Wlemath	Warm estimate of math from PISA 200 study
Female	Gender dummy:1- female;0- male

Medu	Mother tertiary education dummy:1-tertiary education;0- no tertiary education
Fedu	Father tertiary education dummy;1-tertiary education;0- no tertiary education
Esteem	Self-esteem factor
Cope	Coping ability factor
Career	Career orientation factor
Social	Social orientation factor

Table 12

```
Iteration 0:   log pseudolikelihood = -683.30129
Iteration 1:   log pseudolikelihood = -659.99378
Iteration 2:   log pseudolikelihood = -659.94379
Iteration 3:   log pseudolikelihood = -659.94378

Logistic regression                               Number of obs     =      1,001
                                                  Wald chi2(1)      =      44.94
                                                  Prob > chi2       =     0.0000
Log pseudolikelihood = -659.94378                 Pseudo R2         =     0.0342
```

edul	Odds Ratio	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
female	2.47332	.3341136	6.70	0.000	1.89799	3.223047
_cons	.4329897	.0461988	-7.85	0.000	.3512826	.5337015

Note: _cons estimates baseline odds.

```
Marginal effects after logit
      y  = Pr(edul) (predict)
         =  .42342675
```

variable	dy/dx	Std. Err.	z	P>|z|	[95%	C.I.]	X
female*	.214965	.03056	7.03	0.000	.15506	.27487	.583417

(*) dy/dx is for discrete change of dummy variable from 0 to 1

Table 13

```
Iteration 0:   log pseudolikelihood = -666.12023
Iteration 1:   log pseudolikelihood = -555.13903
Iteration 2:   log pseudolikelihood = -553.93517
Iteration 3:   log pseudolikelihood = -553.93421
Iteration 4:   log pseudolikelihood = -553.93421

Logistic regression                               Number of obs     =        973
                                                  Wald chi2(3)      =     149.21
                                                  Prob > chi2       =     0.0000
Log pseudolikelihood = -553.93421                 Pseudo R2         =     0.1684
```

edul	Odds Ratio	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
wlemath	1.003889	.0011902	3.27	0.001	1.001559	1.006224
wleread	1.009049	.0012293	7.39	0.000	1.006643	1.011462
female	2.387162	.3882379	5.35	0.000	1.735586	3.283354
_cons	.0004358	.0002901	-11.63	0.000	.0001182	.0016064

Note: _cons estimates baseline odds.

.

Table 14

```
Marginal effects after logit
      y  = Pr(edul) (predict)
         =  .41177269
```

variable	dy/dx	Std. Err.	z	P>|z|	[95%	C.I.]	X
wlemath	.0009401	.00029	3.29	0.001	.000379	.001501	544.819
wleread	.0021821	.00029	7.46	0.000	.001608	.002756	528.088
female*	.2049238	.0365	5.61	0.000	.13338	.276468	.585817

(*) dy/dx is for discrete change of dummy variable from 0 to 1

Table 15

```
Iteration 0:   log pseudolikelihood = -517.19418
Iteration 1:   log pseudolikelihood = -417.01239
Iteration 2:   log pseudolikelihood = -416.41848
Iteration 3:   log pseudolikelihood = -416.41811
Iteration 4:   log pseudolikelihood = -416.41811

Logistic regression                             Number of obs     =        752
                                                Wald chi2(7)      =     138.26
                                                Prob > chi2       =     0.0000
Log pseudolikelihood = -416.41811               Pseudo R2         =     0.1949
```

edul	Odds Ratio	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
wlemath	1.004334	.0013838	3.14	0.002	1.001626	1.00705
wleread	1.008711	.0013599	6.43	0.000	1.006049	1.01138
female	1.866627	.3775377	3.09	0.002	1.25573	2.774717
esteem	1.052891	.1020657	0.53	0.595	.870701	1.273203
social	1.425416	.1449471	3.49	0.000	1.167844	1.739796
career	.6054343	.0709047	-4.28	0.000	.4812598	.7616483
cope	1.193947	.1295913	1.63	0.102	.9651519	1.476981
_cons	.0004622	.0003625	-9.79	0.000	.0000994	.0021497

Note: _cons estimates baseline odds.

.

Table 16

```
Marginal effects after logit
      y  = Pr(edul) (predict)
         =   .4295299
```

variable	dy/dx	Std. Err.	z	P>|z|	[95%	C.I.]	X
wlemath	.0010597	.00034	3.15	0.002	.000401	.001718	548.694
wleread	.0021253	.00033	6.45	0.000	.001479	.002772	535.366
female*	.1501532	.04725	3.18	0.001	.057553	.242753	.611702
esteem	.0126289	.02375	0.53	0.595	-.033925	.059183	-.005404
social	.0868556	.0249	3.49	0.000	.038045	.135666	-.007502
career	-.1229603	.02872	-4.28	0.000	-.179244	-.066676	-.000099
cope	.043436	.02662	1.63	0.103	-.008745	.095617	.002413

(*) dy/dx is for discrete change of dummy variable from 0 to 1

Table 17

```
Iteration 0:   log pseudolikelihood = -516.59918
Iteration 1:   log pseudolikelihood = -388.61072
Iteration 2:   log pseudolikelihood = -387.83356
Iteration 3:   log pseudolikelihood = -387.81852
Iteration 4:   log pseudolikelihood = -387.81743
Iteration 5:   log pseudolikelihood = -387.81742

Logistic regression                               Number of obs     =        751
                                                  Wald chi2(12)     =     214.57
                                                  Prob > chi2       =     0.0000
Log pseudolikelihood = -387.81742                 Pseudo R2         =     0.2493
```

edul	Odds Ratio	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
medu	2.300165	.5305911	3.61	0.000	1.463551	3.615015
fedu	2.222449	.4310045	4.12	0.000	1.519695	3.250179
urban	1.764987	.3263465	3.07	0.002	1.22844	2.535881
wlemath	1.004229	.001446	2.93	0.003	1.001399	1.007067
wleread	1.009072	.0014151	6.44	0.000	1.006302	1.01185
Age	1.000715	.0001347	5.31	0.000	1.000451	1.000979
female	2.076727	.4369339	3.47	0.001	1.374961	3.136667
wealth	1.032288	.1322865	0.25	0.804	.8030087	1.327032
esteem	1.077871	.1109065	0.73	0.466	.8810149	1.318713
social	1.418572	.1563215	3.17	0.002	1.143015	1.760559
career	.6012508	.0757925	-4.04	0.000	.4696288	.7697621
cope	1.277164	.1483569	2.11	0.035	1.017115	1.603701
_cons	.0001709	.0001436	-10.32	0.000	.0000329	.000887

Note: _cons estimates baseline odds.

Table 18

```
Marginal effects after logit
      y  = Pr(edul) (predict)
         =  .42485041
```

variable	dy/dx	Std. Err.	z	P>|z|	[95%	C.I.]	X
medu*	.2052297	.05546	3.70	0.000	.096527	.313933	.177097
fedu*	.1957217	.04686	4.18	0.000	.103878	.287565	.330226
urban*	.1363318	.0432	3.16	0.002	.051652	.221011	.627164
wlemath	.0010311	.00035	2.94	0.003	.000345	.001717	548.638
wleread	.0022068	.00034	6.47	0.000	.001539	.002875	535.58
Age	.0001747	.00003	5.29	0.000	.00011	.000239	7.13449
female*	.1743606	.04821	3.62	0.000	.079876	.268845	.612517
wealth	.0077649	.03131	0.25	0.804	-.053611	.069141	.05241
esteem	.0183235	.02515	0.73	0.466	-.030978	.067625	-.005864
social	.085438	.02691	3.17	0.001	.032696	.13818	-.007365
career	-.1243127	.03087	-4.03	0.000	-.184807	-.063818	.000526
cope	.059779	.02844	2.10	0.036	.004043	.115515	.002168

(*) dy/dx is for discrete change of dummy variable from 0 to 1

Table 19

Logistic model for edul

Classified	True D	True ~D	Total
+	229	87	316
-	108	327	435
Total	337	414	751

Classified + if predicted Pr(D) >= .5
True D defined as edul != 0

Sensitivity	Pr(+| D)	67.95%
Specificity	Pr(-|~D)	78.99%
Positive predictive value	Pr(D| +)	72.47%
Negative predictive value	Pr(~D| -)	75.17%
False + rate for true ~D	Pr(+|~D)	21.01%
False - rate for true D	Pr(-| D)	32.05%
False + rate for classified +	Pr(~D| +)	27.53%
False - rate for classified -	Pr(D| -)	24.83%
Correctly classified		74.03%

Table 20

```
-> sex = female

note: female omitted because of collinearity
Iteration 0:   log pseudolikelihood = -317.86875
Iteration 1:   log pseudolikelihood =  -250.9352
Iteration 2:   log pseudolikelihood = -250.68049
Iteration 3:   log pseudolikelihood = -250.66542
Iteration 4:   log pseudolikelihood = -250.66514
Iteration 5:   log pseudolikelihood = -250.66514

Logistic regression                             Number of obs     =        460
                                                Wald chi2(11)     =     131.64
                                                Prob > chi2       =     0.0000
Log pseudolikelihood = -250.66514               Pseudo R2         =     0.2114
```

edul	Odds Ratio	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
medu	2.593393	.7695385	3.21	0.001	1.449742	4.639228
fedu	2.314973	.5719919	3.40	0.001	1.426356	3.757196
urban	1.654101	.3852325	2.16	0.031	1.047903	2.610976
wlemath	1.005202	.0019056	2.74	0.006	1.001474	1.008944
wleread	1.008571	.0016832	5.11	0.000	1.005278	1.011876
Age	1.000789	.0001399	5.64	0.000	1.000515	1.001063
female	1	(omitted)				
wealth	.9768299	.1555424	-0.15	0.883	.714958	1.334619
esteem	1.074616	.1367904	0.57	0.572	.8373395	1.379128
social	1.545343	.2449577	2.75	0.006	1.132655	2.108394
career	.6771348	.1083648	-2.44	0.015	.4948274	.9266092
cope	1.171589	.169534	1.09	0.274	.8822727	1.555779
_cons	.0002828	.0002948	-7.84	0.000	.0000367	.0021812

Note: _cons estimates baseline odds.

```
-> sex = male

note: female omitted because of collinearity
Iteration 0:   log pseudolikelihood = -181.56486
Iteration 1:   log pseudolikelihood = -133.96811
Iteration 2:   log pseudolikelihood = -131.21169
Iteration 3:   log pseudolikelihood = -131.18458
Iteration 4:   log pseudolikelihood = -131.18456
Iteration 5:   log pseudolikelihood = -131.18456

Logistic regression                             Number of obs     =        291
                                                Wald chi2(11)     =      74.58
                                                Prob > chi2       =     0.0000
Log pseudolikelihood = -131.18456               Pseudo R2         =     0.2775
```

edul	Odds Ratio	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
medu	1.723367	.6969702	1.35	0.178	.7800668	3.807358
fedu	2.253798	.7397856	2.48	0.013	1.18445	4.288575
urban	1.897502	.6250684	1.94	0.052	.9949065	3.618945
wlemath	1.002501	.0022332	1.12	0.262	.9981337	1.006888
wleread	1.010846	.0027118	4.02	0.000	1.005545	1.016175
Age	.5749784	.1775649	-1.79	0.073	.3138943	1.053221
female	1	(omitted)				
wealth	1.050364	.2402484	0.21	0.830	.6708813	1.644501
esteem	1.048018	.1878749	0.26	0.794	.7375245	1.489227
social	1.276875	.2083628	1.50	0.134	.9273586	1.758122
career	.4570201	.1012553	-3.53	0.000	.2960376	.7055434
cope	1.471104	.2830136	2.01	0.045	1.008992	2.144861
_cons	3.583049	19.29034	0.24	0.813	.0000937	137067.2

Note: _cons estimates baseline odds.

Table 21 2007

```
Iteration 0:   log pseudolikelihood = -2008.1348
Iteration 1:   log pseudolikelihood = -2008.0892
Iteration 2:   log pseudolikelihood = -2008.0892

Logistic regression                             Number of obs     =      2,902
                                                Wald chi2(1)      =       0.09
                                                Prob > chi2       =     0.7626
Log pseudolikelihood = -2008.0892               Pseudo R2         =     0.0000
```

edul	Odds Ratio	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
female	.9773424	.0741674	-0.30	0.763	.8422709	1.134075
_cons	.9205298	.054111	-1.41	0.159	.8203555	1.032936

Note: _cons estimates baseline odds.

```
Marginal effects after logit
      y  = Pr(edul) (predict)
         =  .47587796
```

variable	dy/dx	Std. Err.	z	P>|z|	[95%	C.I.]	X
female*	-.0057168	.01893	-0.30	0.763	-.042821	.031387	.600276

(*) dy/dx is for discrete change of dummy variable from 0 to 1

Table 22 2007

```
Iteration 0:   log pseudolikelihood = -394.30405
Iteration 1:   log pseudolikelihood = -317.75165
Iteration 2:   log pseudolikelihood = -317.55597
Iteration 3:   log pseudolikelihood = -317.55593
Iteration 4:   log pseudolikelihood = -317.55593

Logistic regression                               Number of obs     =        570
                                                  Wald chi2(3)      =     108.16
                                                  Prob > chi2       =     0.0000
Log pseudolikelihood = -317.55593                 Pseudo R2         =     0.1946
```

edul	Odds Ratio	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
wlemath	1.00625	.0017155	3.65	0.000	1.002893	1.009618
wleread	1.010666	.0016722	6.41	0.000	1.007394	1.013949
female	.981062	.2155559	-0.09	0.931	.6377833	1.509106
_cons	.0000782	.0000738	-10.02	0.000	.0000123	.0004974

Note: _cons estimates baseline odds.

```
Marginal effects after logit
      y  = Pr(edul) (predict)
         =  .46051745
```

variable	dy/dx	Std. Err.	z	P>|z|	[95%	C.I.]	X
wlemath	.0015479	.00042	3.65	0.000	.000718	.002378	561.855
wleread	.0026358	.00041	6.43	0.000	.001833	.003439	547.602
female*	-.0047509	.05461	-0.09	0.931	-.111777	.102275	.615789

(*) dy/dx is for discrete change of dummy variable from 0 to 1

Table 23 2007

```
Iteration 0:   log pseudolikelihood = -391.52588
Iteration 1:   log pseudolikelihood = -312.25326
Iteration 2:   log pseudolikelihood = -312.01328
Iteration 3:   log pseudolikelihood = -312.01323
Iteration 4:   log pseudolikelihood = -312.01323

Logistic regression                               Number of obs     =        566
                                                  Wald chi2(7)      =     112.19
                                                  Prob > chi2       =     0.0000
Log pseudolikelihood = -312.01323                 Pseudo R2         =     0.2031
```

edul	Odds Ratio	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
wlemath	1.006554	.0017741	3.71	0.000	1.003083	1.010037
wleread	1.010625	.0016582	6.44	0.000	1.00738	1.01388
female	.7483813	.1902531	-1.14	0.254	.4547063	1.231728
Esteem	1.041649	.1280008	0.33	0.740	.8186971	1.325316
Social	1.234572	.1658898	1.57	0.117	.9487249	1.606544
Career	.8496412	.1071212	-1.29	0.196	.6636172	1.087811
Cope	1.283382	.1617592	1.98	0.048	1.002465	1.643019
_cons	.000078	.000075	-9.84	0.000	.0000119	.0005138

Note: _cons estimates baseline odds.

```
Marginal effects after logit
      y  = Pr(edul) (predict)
         =  .45849118
```

variable	dy/dx	Std. Err.	z	P>|z|	[95% C.I.]		X
wlemath	.0016219	.00044	3.71	0.000	.000764	.00248	562.056
wleread	.002624	.00041	6.47	0.000	.001829	.003419	547.913
female*	-.0720187	.0631	-1.14	0.254	-.195701	.051664	.616608
Esteem	.0101309	.03051	0.33	0.740	-.049664	.069925	.02583
Social	.052318	.03335	1.57	0.117	-.013047	.117683	-.01729
Career	-.0404545	.03129	-1.29	0.196	-.101778	.020869	-.026298
Cope	.0619448	.0313	1.98	0.048	.000594	.123296	.025585

(*) dy/dx is for discrete change of dummy variable from 0 to 1

Table 24 2007

```
Iteration 0:   log pseudolikelihood = -390.77729
Iteration 1:   log pseudolikelihood = -283.37509
Iteration 2:   log pseudolikelihood = -282.48631
Iteration 3:   log pseudolikelihood = -282.47146
Iteration 4:   log pseudolikelihood = -282.46938
Iteration 5:   log pseudolikelihood = -282.46932
Iteration 6:   log pseudolikelihood = -282.46932

Logistic regression                             Number of obs     =        565
                                                Wald chi2(12)     =     155.59
                                                Prob > chi2       =     0.0000
Log pseudolikelihood = -282.46932               Pseudo R2         =     0.2772
```

edul	Odds Ratio	Robust Std. Err.	z	P>|z|	[95% Conf. Interval]	
medu	2.152812	.6103863	2.70	0.007	1.234994	3.752728
fedu	2.444281	.5752814	3.80	0.000	1.54104	3.876933
urban	1.167268	.2568733	0.70	0.482	.7583191	1.796756
wlemath	1.006947	.0019341	3.60	0.000	1.003163	1.010745
wleread	1.011706	.0017953	6.56	0.000	1.008193	1.015231
Age	1.000751	.0001487	5.06	0.000	1.00046	1.001043
female	.7861635	.2078632	-0.91	0.363	.4682238	1.319995
wealth	1.649443	.2744667	3.01	0.003	1.190413	2.285477
Esteem	1.000932	.1348166	0.01	0.994	.7686977	1.303329
Social	1.206935	.1722615	1.32	0.188	.9124207	1.596515
Career	.8387693	.1152254	-1.28	0.201	.6407805	1.097933
Cope	1.407502	.1885157	2.55	0.011	1.082535	1.83002
_cons	.0000185	.0000204	-9.89	0.000	2.14e-06	.0001603

Note: _cons estimates baseline odds.

Table 25 2007

```
Marginal effects after logit
      y  = Pr(edul) (predict)
         =  .45449116
```

variable	dy/dx	Std. Err.	z	P>|z|	[95% C.I.	]	X
medu*	.1892096	.06779	2.79	0.005	.056337	.322082	.182301
fedu*	.2196753	.05619	3.91	0.000	.109544	.329807	.348673
urban*	.03826	.05425	0.71	0.481	-.068073	.144593	.619469
wlemath	.0017164	.00048	3.61	0.000	.000785	.002648	562.004
wleread	.0028853	.00044	6.55	0.000	.002021	.003749	548.219
Age	.0001862	.00004	5.04	0.000	.000114	.000259	3.61947
female*	-.0597207	.06567	-0.91	0.363	-.188435	.068994	.617699
wealth	.1240729	.0413	3.00	0.003	.043131	.205015	.062407
Esteem	.0002311	.03339	0.01	0.994	-.06522	.065682	.022869
Social	.0466316	.03536	1.32	0.187	-.022666	.115929	-.016062
Career	-.0435908	.03406	-1.28	0.201	-.110343	.023162	-.027693
Cope	.0847461	.03323	2.55	0.011	.019625	.149867	.026455

(*) dy/dx is for discrete change of dummy variable from 0 to 1

Table 26 2007

Logistic model for edul

Classified	True D	True ~D	Total
+	191	65	256
-	76	233	309
Total	267	298	565

Classified + if predicted Pr(D) >= .5
True D defined as edul != 0

Sensitivity	Pr(+| D)	71.54%
Specificity	Pr(-|~D)	78.19%
Positive predictive value	Pr(D| +)	74.61%
Negative predictive value	Pr(~D| -)	75.40%
False + rate for true ~D	Pr(+|~D)	21.81%
False - rate for true D	Pr(-| D)	28.46%
False + rate for classified +	Pr(~D| +)	25.39%
False - rate for classified -	Pr(D| -)	24.60%
Correctly classified		75.04%

.

Table 27 2007

```
-> sex = female

note: female omitted because of collinearity
Iteration 0:   log pseudolikelihood = -241.79231
Iteration 1:   log pseudolikelihood =  -166.4027
Iteration 2:   log pseudolikelihood =  -165.1302
Iteration 3:   log pseudolikelihood =  -165.1034
Iteration 4:   log pseudolikelihood =   -165.102
Iteration 5:   log pseudolikelihood = -165.10192
Iteration 6:   log pseudolikelihood = -165.10192

Logistic regression                             Number of obs     =        349
                                                Wald chi2(11)     =     110.52
                                                Prob > chi2       =     0.0000
Log pseudolikelihood = -165.10192               Pseudo R2         =     0.3172

------------------------------------------------------------------------------
             |               Robust
        edul | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        medu |   3.277834    1.27826     3.04   0.002     1.526311    7.039322
        fedu |   2.337397   .7268034     2.73   0.006     1.270735     4.29942
       urban |    1.30367   .3687602     0.94   0.349     .7488485    2.269557
     wlemath |   1.004908   .0023805     2.07   0.039     1.000253    1.009584
     wleread |   1.016249   .0025716     6.37   0.000     1.011222    1.021302
         Age |   1.000746   .0001522     4.90   0.000     1.000448    1.001045
      female |          1  (omitted)
      wealth |   2.407187   .5160293     4.10   0.000     1.581385    3.664225
      Esteem |   1.238662   .2374496     1.12   0.264     .8507046    1.803545
      Social |   1.150451   .2184201     0.74   0.460     .7929805    1.669068
      Career |   .7711351   .1407813    -1.42   0.155     .5391777    1.102882
        Cope |   1.516995   .2788757     2.27   0.023     1.058045    2.175025
       _cons |   3.06e-06   4.94e-06    -7.86   0.000     1.29e-07    .0000726
------------------------------------------------------------------------------
Note: _cons estimates baseline odds.

-------------------------------------------------------------------------------------------------------------

-> sex = male

note: female omitted because of collinearity
Iteration 0:   log pseudolikelihood = -148.59748
Iteration 1:   log pseudolikelihood = -108.00417
Iteration 2:   log pseudolikelihood = -107.48211
Iteration 3:   log pseudolikelihood = -107.48002
Iteration 4:   log pseudolikelihood = -107.48002

Logistic regression                             Number of obs     =        216
                                                Wald chi2(11)     =      54.06
                                                Prob > chi2       =     0.0000
Log pseudolikelihood = -107.48002               Pseudo R2         =     0.2767

------------------------------------------------------------------------------
             |               Robust
        edul | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        medu |    1.69379   .8142507     1.10   0.273     .6601821    4.345655
        fedu |   2.596479   .9646965     2.57   0.010     1.253514    5.378243
       urban |   .7221555    .266454    -0.88   0.378     .3503987    1.488329
     wlemath |   1.010126   .0031288     3.25   0.001     1.004013    1.016277
     wleread |   1.007442   .0026721     2.80   0.005     1.002219    1.012693
         Age |   .7904727   .2262919    -0.82   0.411     .4510346    1.385364
      female |          1  (omitted)
      wealth |   1.061069   .3381628     0.19   0.852     .5681534    1.981626
      Esteem |   .7265245   .1556789    -1.49   0.136     .4773706    1.105719
      Social |   1.241605    .281104     0.96   0.339     .7966517    1.935079
      Career |   .9758435   .2413019    -0.10   0.921     .6010342    1.584387
        Cope |   1.331469   .3019755     1.26   0.207     .8536495    2.076742
       _cons |   .0027649   .0146264    -1.11   0.265     8.69e-08     88.0017
------------------------------------------------------------------------------
Note: _cons estimates baseline odds.
```

Table 28 2007

jobfield	Freq.	Percent	Cum.
Physical,math,engineering prof.and asso	89	8.44	8.44
Teaching prof. and associates	57	5.40	13.84
Social sciences,Business,law prof.and a	156	14.79	28.63
Life and health science prof. and assoc	89	8.44	37.06
clerks/office clerks	240	22.75	59.81
service workers	177	16.78	76.59
agriculture	33	3.13	79.72
craft, plant and machine operators	214	20.28	100.00
Total	1,055	100.00	

Table 29 2007

-> sex = female

jobfield	Freq.	Percent	Cum.
Physical,math,engineering prof.and asso	29	4.46	4.46
Teaching prof. and associates	53	8.15	12.62
Social sciences,Business,law prof.and a	106	16.31	28.92
Life and health science prof. and assoc	86	13.23	42.15
clerks/office clerks	184	28.31	70.46
service workers	129	19.85	90.31
agriculture	18	2.77	93.08
craft, plant and machine operators	45	6.92	100.00
Total	650	100.00	

-> sex = male

jobfield	Freq.	Percent	Cum.
Physical,math,engineering prof.and asso	60	14.81	14.81
Teaching prof. and associates	4	0.99	15.80
Social sciences,Business,law prof.and a	50	12.35	28.15
Life and health science prof. and assoc	3	0.74	28.89
clerks/office clerks	56	13.83	42.72
service workers	48	11.85	54.57
agriculture	15	3.70	58.27
craft, plant and machine operators	169	41.73	100.00
Total	405	100.00	

Table 30 Multinomial logit analysis of job field with respect to Female dummy variable and cognitive abilities

```
Iteration 0:   log pseudolikelihood = -406.70086
Iteration 1:   log pseudolikelihood = -342.14466
Iteration 2:   log pseudolikelihood = -335.02647
Iteration 3:   log pseudolikelihood = -334.59233
Iteration 4:   log pseudolikelihood = -334.50582
Iteration 5:   log pseudolikelihood = -334.48442
Iteration 6:   log pseudolikelihood = -334.48001
Iteration 7:   log pseudolikelihood = -334.47906
Iteration 8:   log pseudolikelihood = -334.47885
Iteration 9:   log pseudolikelihood =  -334.4788
Iteration 10:  log pseudolikelihood = -334.47879

Multinomial logistic regression                 Number of obs     =        216
                                                Wald chi2(21)     =    2249.03
                                                Prob > chi2       =     0.0000
Log pseudolikelihood = -334.47879               Pseudo R2         =     0.1776
```

jobfield	RRR	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
Physical_math_engineering_prof_a						
female	.2424851	.1381043	-2.49	0.013	.0794135	.7404164
wlemath	1.007038	.0040168	1.76	0.079	.9991959	1.014942
wleread	1.005821	.0039838	1.47	0.143	.9980435	1.01366
_cons	.000499	.0011314	-3.35	0.001	5.87e-06	.0424607
Teaching_prof__and_associates						
female	2.236776	2.003199	0.90	0.369	.3866415	12.94007
wlemath	.993707	.0054025	-1.16	0.246	.9831746	1.004352
wleread	1.015869	.0051926	3.08	0.002	1.005743	1.026098
_cons	.0005407	.0020354	-2.00	0.046	3.38e-07	.8654917
Social_sciences_Business_law_pro						
female	.7788046	.4174708	-0.47	0.641	.2723642	2.226932
wlemath	.9997171	.0040335	-0.07	0.944	.9918428	1.007654
wleread	1.004458	.0038994	1.15	0.252	.9968444	1.01213
_cons	.0460329	.0948529	-1.49	0.135	.0008112	2.612213
Life_and_health_science_prof__an						
female	2690198	1109078	35.91	0.000	1199120	6035396
wlemath	1.001955	.004695	0.42	0.677	.9927954	1.0112
wleread	1.003508	.0040598	0.87	0.387	.9955823	1.011497
_cons	4.96e-09	1.53e-08	-6.18	0.000	1.16e-11	2.13e-06
clerks_office_clerks	(base outcome)					
service_workers						
female	2.362775	1.476343	1.38	0.169	.6943244	8.040488
wlemath	.9962503	.0036596	-1.02	0.306	.9891034	1.003449
wleread	.9866748	.0042385	-3.12	0.002	.9784023	.9950172
_cons	1455.795	2694.716	3.93	0.000	38.68147	54789.5
agriculture						
female	.3226588	.2236855	-1.63	0.103	.0829167	1.255582
wlemath	.9984722	.004371	-0.35	0.727	.9899418	1.007076
wleread	.9973285	.0072136	-0.37	0.711	.9832899	1.011568
_cons	1.877785	7.115153	0.17	0.868	.0011177	3154.647
craft__plant__and_machine_operat						
female	.0929477	.0529391	-4.17	0.000	.030439	.2838226
wlemath	1.008588	.0037792	2.28	0.022	1.001208	1.016023
wleread	.9876118	.0040014	-3.08	0.002	.9798002	.9954857
_cons	7.34959	12.56121	1.17	0.243	.2579121	209.4376

Note: _cons estimates baseline relative risk for each outcome.

.

Table 31

```
. mfx, predict(pr outcome(1))

Marginal effects after mlogit
      y  = Pr(jobfield==Physical_math_engineering_prof_a) (predict, pr outcome(1))
         =  .08379177
```

variable	dy/dx	Std. Err.	z	P>|z|	[95% C.I.]		X
female*	-.1035328	.01943	-5.33	0.000	-.141616	-.06545	.616114

```
(*) dy/dx is for discrete change of dummy variable from 0 to 1
```

Table 32

```
. mfx, predict(pr outcome(2))

Marginal effects after mlogit
      y  = Pr(jobfield==Teaching_prof__and_associates) (predict, pr outcome(2))
         =  .04295984
```

variable	dy/dx	Std. Err.	z	P>|z|	[95% C.I.]		X
female*	.0716619	.01181	6.07	0.000	.048513	.094811	.616114

```
(*) dy/dx is for discrete change of dummy variable from 0 to 1
```

Table 33

```
. mfx, predict(pr outcome(3))

Marginal effects after mlogit
      y  = Pr(jobfield==Social_sciences_Business_law_pro) (predict, pr outcome(3))
         =  .17362811
```

variable	dy/dx	Std. Err.	z	P>|z|	[95% C.I.]		X
female*	.0396201	.02185	1.81	0.070	-.003214	.082454	.616114

```
(*) dy/dx is for discrete change of dummy variable from 0 to 1
```

Table 34

```
. mfx, predict(pr outcome(4))

Marginal effects after mlogit
      y  = Pr(jobfield==Life_and_health_science_prof__an) (predict, pr outcome(4))
         =   .0518348
```

variable	dy/dx	Std. Err.	z	P>|z|	[95% C.I.]		X
female*	.1249003	.01396	8.95	0.000	.097534	.152267	.616114

```
(*) dy/dx is for discrete change of dummy variable from 0 to 1
```

Table 35

```
Iteration 0:   log pseudolikelihood = -406.70086
Iteration 1:   log pseudolikelihood = -342.14466
Iteration 2:   log pseudolikelihood = -335.02647
Iteration 3:   log pseudolikelihood = -334.59233
Iteration 4:   log pseudolikelihood = -334.50582
Iteration 5:   log pseudolikelihood = -334.48442
Iteration 6:   log pseudolikelihood = -334.48001
Iteration 7:   log pseudolikelihood = -334.47906
Iteration 8:   log pseudolikelihood = -334.47885
Iteration 9:   log pseudolikelihood =  -334.4788
Iteration 10:  log pseudolikelihood = -334.47879

Multinomial logistic regression                 Number of obs     =        216
                                                Wald chi2(21)     =    2249.03
                                                Prob > chi2       =     0.0000
Log pseudolikelihood = -334.47879               Pseudo R2         =     0.1776
```

jobfield	RRR	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
Physical_math_engineering_prof_a						
female	.2424851	.1381043	-2.49	0.013	.0794135	.7404164
wlemath	1.007038	.0040168	1.76	0.079	.9991959	1.014942
wleread	1.005821	.0039838	1.47	0.143	.9980435	1.01366
_cons	.000499	.0011314	-3.35	0.001	5.87e-06	.0424607
Teaching_prof__and_associates						
female	2.236776	2.003199	0.90	0.369	.3866415	12.94007
wlemath	.993707	.0054025	-1.16	0.246	.9831746	1.004352
wleread	1.015869	.0051926	3.08	0.002	1.005743	1.026098
_cons	.0005407	.0020354	-2.00	0.046	3.38e-07	.8654917
Social_sciences_Business_law_pro						
female	.7788046	.4174708	-0.47	0.641	.2723642	2.226932
wlemath	.9997171	.0040335	-0.07	0.944	.9918428	1.007654
wleread	1.004458	.0038994	1.15	0.252	.9968444	1.01213
_cons	.0460329	.0948529	-1.49	0.135	.0008112	2.612213
Life_and_health_science_prof__an						
female	2690198	1109078	35.91	0.000	1199120	6035396
wlemath	1.001955	.004695	0.42	0.677	.9927954	1.0112
wleread	1.003508	.0040598	0.87	0.387	.9955823	1.011497
_cons	4.96e-09	1.53e-08	-6.18	0.000	1.16e-11	2.13e-06
clerks_office_clerks	(base outcome)					
service_workers						
female	2.362775	1.476343	1.38	0.169	.6943244	8.040488
wlemath	.9962503	.0036596	-1.02	0.306	.9891034	1.003449
wleread	.9866748	.0042385	-3.12	0.002	.9784023	.9950172
_cons	1455.795	2694.716	3.93	0.000	38.68147	54789.5
agriculture						
female	.3226588	.2236855	-1.63	0.103	.0829167	1.255582
wlemath	.9984722	.004371	-0.35	0.727	.9899418	1.007076
wleread	.9973285	.0072136	-0.37	0.711	.9832899	1.011568
_cons	1.877785	7.115153	0.17	0.868	.0011177	3154.647
craft__plant__and_machine_operat						
female	.0929477	.0529391	-4.17	0.000	.030439	.2838226
wlemath	1.008588	.0037792	2.28	0.022	1.001208	1.016023
wleread	.9876118	.0040014	-3.08	0.002	.9798002	.9954857
_cons	7.34959	12.56121	1.17	0.243	.2579121	209.4376

Note: _cons estimates baseline relative risk for each outcome.

.

Table 36

```
Marginal effects after mlogit
      y  = Pr(jobfield==Physical_math_engineering_prof_a) (predict, pr outcome(1))
         =  .08289774
```

variable	dy/dx	Std. Err.	z	P>|z|	[95%	C.I.]	X
female*	-.0841803	.04132	-2.04	0.042	-.165174	-.003187	.592593
wlemath	.0005191	.00028	1.89	0.059	-.00002	.001058	540.144
wleread	.0006114	.00028	2.16	0.031	.000056	.001167	518.53

(*) dy/dx is for discrete change of dummy variable from 0 to 1

Table 37

```
. mfx, predict(pr outcome(2))

Marginal effects after mlogit
      y  = Pr(jobfield==Teaching_prof__and_associates) (predict, pr outcome(2))
         =  .04340534
```

variable	dy/dx	Std. Err.	z	P>|z|	[95%	C.I.]	X
female*	.0403653	.02715	1.49	0.137	-.012847	.093577	.592593
wlemath	-.0003066	.00026	-1.16	0.246	-.000825	.000212	540.144
wleread	.0007516	.00025	2.97	0.003	.000255	.001248	518.53

(*) dy/dx is for discrete change of dummy variable from 0 to 1

Table 38

```
. mfx, predict(pr outcome(3))

Marginal effects after mlogit
      y  = Pr(jobfield==Social_sciences_Business_law_pro) (predict, pr outcome(3))
         =  .14681251
```

variable	dy/dx	Std. Err.	z	P>|z|	[95%	C.I.]	X
female*	.0140215	.05188	0.27	0.787	-.087653	.115696	.592593
wlemath	-.0001518	.00047	-0.33	0.745	-.001066	.000762	540.144
wleread	.0008837	.00043	2.06	0.040	.000042	.001726	518.53

(*) dy/dx is for discrete change of dummy variable from 0 to 1

Table 39

```
. mfx, predict(pr outcome(4))

Marginal effects after mlogit
      y  = Pr(jobfield==Life_and_health_science_prof__an) (predict, pr outcome(4))
         =  .00024297
```

variable	dy/dx	Std. Err.	z	P>|z|	[95%	C.I.]	X
female*	.0982738	.02916	3.37	0.001	.041125	.155423	.592593
wlemath	2.92e-07	.00000	0.28	0.778	-1.7e-06	2.3e-06	540.144
wleread	1.23e-06	.00000	1.49	0.137	-3.9e-07	2.9e-06	518.53

(*) dy/dx is for discrete change of dummy variable from 0 to 1

Table 40 Multinomial logit analysis of education path with respect to four factors, Female dummy variable and cognitive abilities

Iteration 0: log pseudolikelihood = -404.83941
Iteration 1: log pseudolikelihood = -319.23048
Iteration 2: log pseudolikelihood = -308.84874
Iteration 3: log pseudolikelihood = -307.99415
Iteration 4: log pseudolikelihood = -307.88146
Iteration 5: log pseudolikelihood = -307.86328
Iteration 6: log pseudolikelihood = -307.85886
Iteration 7: log pseudolikelihood = -307.85793
Iteration 8: log pseudolikelihood = -307.85773
Iteration 9: log pseudolikelihood = -307.85769
Iteration 10: log pseudolikelihood = -307.85768

Multinomial logistic regression

Number of obs	=	215
Wald chi2(49)	=	1982.49
Prob > chi2	=	0.0000
Pseudo R2	=	0.2396

Log pseudolikelihood = -307.85768

jobfield	RRR	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
Physical_math_engineering_prof_a	(base outcome)					
Teaching_prof__and_associates						
wleread	1.013004	.0065283	2.00	0.045	1.000289	1.02588
wlemath	.9882223	.0070828	-1.65	0.098	.9744373	1.002202
female	3.030544	3.400929	0.99	0.323	.33596	27.33718
Esteem	.4539959	.2070091	-1.73	0.083	.1857498	1.109623
Social	5.724399	3.670237	2.72	0.007	1.629207	20.11331
Career	.7775983	.4013835	-0.49	0.626	.2827347	2.13861
Cope	2.37574	1.41148	1.46	0.145	.7414483	7.612316
_cons	.1173762	.6040699	-0.42	0.677	4.89e-06	2819.947
Social_sciences_Business_law_pro						
wleread	.9989904	.0053765	-0.19	0.851	.9885081	1.009584
wlemath	.9925018	.0052395	-1.43	0.154	.9822855	1.002824
female	2.352656	1.956282	1.03	0.304	.46107	12.00466
Esteem	.8616735	.3271236	-0.39	0.695	.4094442	1.813388
Social	1.285224	.6897422	0.47	0.640	.4489148	3.679543
Career	1.35207	.7698819	0.53	0.596	.4429113	4.127449
Cope	1.443566	.7281949	0.73	0.467	.5371001	3.879879
_cons	127.905	354.3776	1.75	0.080	.5604299	29191.31
Life_and_health_science_prof__an						
wleread	.9982762	.0056514	-0.30	0.761	.9872609	1.009414
wlemath	.9953566	.0066907	-0.69	0.489	.9823291	1.008557
female	3057346	2385434	19.14	0.000	662539	1.41e+07
Esteem	1.007344	.4369253	0.02	0.987	.4305044	2.357099
Social	5.258348	3.667009	2.38	0.017	1.340441	20.6277
Career	1.968654	1.210351	1.10	0.271	.5899849	6.56898
Cope	1.741661	.8646206	1.12	0.264	.6582607	4.60818
_cons	.0000132	.0000532	-2.79	0.005	4.89e-09	.0356967
clerks_office_clerks						
wleread	.9956124	.0045449	-0.96	0.335	.9867444	1.00456
wlemath	.9916883	.0045688	-1.81	0.070	.9827739	1.000684
female	2.78066	1.865598	1.52	0.127	.7465524	10.35703
Esteem	.8205364	.2653028	-0.61	0.541	.4353917	1.546378
Social	1.271135	.5272937	0.58	0.563	.5637609	2.866078
Career	2.836239	1.116992	2.65	0.008	1.310736	6.1372
Cope	2.1668	.9860883	1.70	0.089	.8880674	5.286788
_cons	2976.967	7084.104	3.36	0.001	28.06873	315736.9
service_workers						
wleread	.9810774	.0054251	-3.45	0.001	.9705018	.9917682
wlemath	.9896183	.0056909	-1.81	0.070	.9785269	1.000835
female	3.46644	3.105995	1.39	0.165	.5986753	20.07133
Esteem	1.445857	.6056312	0.88	0.379	.6361792	3.286027
Social	3.40145	1.703779	2.44	0.015	1.27438	9.078814
Career	1.991802	.9102181	1.51	0.132	.8133213	4.877869
Cope	2.37235	1.255899	1.63	0.103	.8405451	6.695708
_cons	4747948	1.30e+07	5.60	0.000	21885.14	1.03e+09
agriculture						
wleread	.9916656	.0079604	-1.04	0.297	.9761856	1.007391
wlemath	.9892909	.0064013	-1.66	0.096	.9768239	1.001917
female	1.905625	1.690246	0.73	0.467	.3349973	10.84011
Esteem	1.104127	.3809126	0.29	0.774	.5615173	2.171075
Social	.6081733	.2857244	-1.06	0.290	.2421735	1.527313
Career	1.047143	.6024075	0.08	0.936	.3390944	3.233635
Cope	1.198381	.5863137	0.37	0.711	.4593462	3.126435
_cons	8346.653	38014.12	1.98	0.047	1.108632	6.28e+07
craft__plant__and_machine_operat						
wleread	.981912	.0051722	-3.47	0.001	.9718268	.9921019
wlemath	1.000944	.004915	0.19	0.848	.9913574	1.010624
female	.1953028	.1549713	-2.06	0.040	.0412369	.9249759
Esteem	1.229241	.4266212	0.59	0.552	.6226138	2.426921
Social	1.787573	.7528303	1.38	0.168	.7830385	4.080792
Career	1.159005	.5094856	0.34	0.737	.4896733	2.743241
Cope	1.836192	.9063321	1.23	0.218	.697864	4.831317
_cons	38516.13	102265.3	3.98	0.000	211.6377	7009586

Note: _cons estimates baseline relative risk for each outcome.

Table 41

```
. mfx, predict(pr)

Marginal effects after mlogit
      y  = Pr(jobfield==Physical_math_engineering_prof_a) (predict, pr)
         =  .07586278
```

variable	dy/dx	Std. Err.	z	P>|z|	[95%	C.I.]	X
wleread	.000529	.00032	1.64	0.101	-.000103	.001161	518.761
wlemath	.0005117	.00028	1.80	0.071	-.000045	.001068	540.482
female*	-.0347195	.04025	-0.86	0.388	-.113608	.04417	.590698
Esteem	.0040119	.02059	0.19	0.846	-.036349	.044373	.011553
Social	-.0296558	.02914	-1.02	0.309	-.086762	.02745	-.052786
Career	-.0471575	.02577	-1.83	0.067	-.09766	.003345	.048372
Cope	-.046966	.0284	-1.65	0.098	-.102632	.0087	-.07255

(*) dy/dx is for discrete change of dummy variable from 0 to 1

Table 42

```
. mfx, predict(pr outcome(2))

Marginal effects after mlogit
      y  = Pr(jobfield==Teaching_prof__and_associates) (predict, pr outcome(2))
         =  .01938231
```

variable	dy/dx	Std. Err.	z	P>|z|	[95%	C.I.]	X
wleread	.0003856	.00023	1.69	0.091	-.000061	.000832	518.761
wlemath	-.0000989	.00015	-0.66	0.509	-.000392	.000195	540.482
female*	.0096741	.01362	0.71	0.477	-.017013	.036361	.590698
Esteem	-.0142806	.00933	-1.53	0.126	-.032562	.004001	.011553
Social	.0262402	.01508	1.74	0.082	-.003309	.05579	-.052786
Career	-.0169239	.0118	-1.43	0.151	-.04005	.006202	.048372
Cope	.0047723	.00922	0.52	0.605	-.013307	.022851	-.07255

(*) dy/dx is for discrete change of dummy variable from 0 to 1

Table 43

```
. mfx, predict(pr outcome(3))

Marginal effects after mlogit
      y  = Pr(jobfield==Social_sciences_Business_law_pro) (predict, pr outcome(3))
         =  .15864906
```

variable	dy/dx	Std. Err.	z	P>|z|	[95%	C.I.]	X
wleread	.000946	.00047	2.00	0.045	.00002	.001872	518.761
wlemath	-.0001239	.00048	-0.26	0.798	-.001074	.000826	540.482
female*	.0466358	.07324	0.64	0.524	-.09691	.190181	.590698
Esteem	-.0152296	.04127	-0.37	0.712	-.096123	.065664	.011553
Social	-.0222079	.04935	-0.45	0.653	-.118929	.074513	-.052786
Career	-.0507644	.06051	-0.84	0.401	-.169356	.067827	.048372
Cope	-.0399756	.04061	-0.98	0.325	-.119577	.039625	-.07255

(*) dy/dx is for discrete change of dummy variable from 0 to 1

Table 44 M.E. on probability of choosing life and health science professions

```
. mfx, predict(pr outcome(4))

Marginal effects after mlogit
      y  = Pr(jobfield==Life_and_health_science_prof__an) (predict, pr outcome(4))
         =  .00020384
```

variable	dy/dx	Std. Err.	z	P>|z|	[95% C.I.]		X
wleread	1.07e-06	.00000	1.17	0.243	-7.3e-07	2.9e-06	518.761
wlemath	4.26e-07	.00000	0.40	0.692	-1.7e-06	2.5e-06	540.482
female*	.064263	.03206	2.00	0.045	.001428	.127098	.590698
Esteem	.0000123	.00007	0.18	0.856	-.000121	.000145	.011553
Social	.0002587	.00008	3.08	0.002	.000094	.000423	-.052786
Career	.0000114	.0001	0.11	0.912	-.000189	.000212	.048372
Cope	-.0000131	.00006	-0.21	0.834	-.000136	.00011	-.07255

(*) dy/dx is for discrete change of dummy variable from 0 to 1

Table 45 M.E. on probability of choosing office clerks professions

```
. mfx, predict(pr outcome(5))

Marginal effects after mlogit
      y  = Pr(jobfield==clerks_office_clerks) (predict, pr outcome(5))
         =  .45434598
```

variable	dy/dx	Std. Err.	z	P>|z|	[95% C.I.]		X
wleread	.0011703	.0007	1.67	0.095	-.000204	.002544	518.761
wlemath	-.0007275	.00066	-1.10	0.270	-.002021	.000566	540.482
female*	.1952496	.0872	2.24	0.025	.02434	.366159	.590698
Esteem	-.0658411	.05306	-1.24	0.215	-.169844	.038161	.011553
Social	-.0686082	.05456	-1.26	0.209	-.175536	.03832	-.052786
Career	.1912173	.06426	2.98	0.003	.065271	.317163	.048372
Cope	.070042	.05567	1.26	0.208	-.03906	.179144	-.07255

(*) dy/dx is for discrete change of dummy variable from 0 to 1

Table 46 . M.E. on probability of choosing service workers professions

```
. mfx, predict(pr outcome(6))

Marginal effects after mlogit
      y  = Pr(jobfield==service_workers) (predict, pr outcome(6))
         =  .12101414
```

variable	dy/dx	Std. Err.	z	P>|z|	[95% C.I.]		X
wleread	-.001468	.00044	-3.37	0.001	-.002321	-.000615	518.761
wlemath	-.0004466	.00043	-1.04	0.300	-.001291	.000398	540.482
female*	.0734727	.0523	1.40	0.160	-.029032	.175978	.590698
Esteem	.0510178	.03588	1.42	0.155	-.019305	.121341	.011553
Social	.1008396	.04046	2.49	0.013	.021544	.180136	-.052786
Career	.0081592	.03226	0.25	0.800	-.05507	.071389	.048372
Cope	.029623	.03498	0.85	0.397	-.038943	.098189	-.07255

(*) dy/dx is for discrete change of dummy variable from 0 to 1

Table 47 M.E. on probability of choosing agriculture professions

```
. mfx, predict(pr outcome(7))

Marginal effects after mlogit
      y  = Pr(jobfield==agriculture) (predict, pr outcome(7))
         =   .0364142
```

variable	dy/dx	Std. Err.	z	P>|z|	[95% C.I.	]	X
wleread	-.0000508	.00025	-0.20	0.839	-.00054	.000438	518.761
wlemath	-.0001464	.00019	-0.76	0.446	-.000523	.00023	540.482
female*	.0043535	.02054	0.21	0.832	-.035913	.04462	.590698
Esteem	.0055327	.00946	0.58	0.559	-.013013	.024079	.011553
Social	-.0323434	.01425	-2.27	0.023	-.060273	-.004414	-.052786
Career	-.0209582	.01893	-1.11	0.268	-.058052	.016136	.048372
Cope	-.0159538	.01454	-1.10	0.273	-.044457	.012549	-.07255

(*) dy/dx is for discrete change of dummy variable from 0 to 1

Table 48 M.E. on probability of choosing craft, plant and machine operating professions

```
Marginal effects after mlogit
      y  = Pr(jobfield==craft__plant__and_machine_operat) (predict, pr outcome(8))
         =  .13412769
```

variable	dy/dx	Std. Err.	z	P>|z|	[95% C.I.	]	X
wleread	-.001513	.00045	-3.37	0.001	-.002393	-.000633	518.761
wlemath	.0010313	.0004	2.59	0.010	.00025	.001812	540.482
female*	-.3589293	.08375	-4.29	0.000	-.523077	-.194782	.590698
Esteem	.0347767	.03006	1.16	0.247	-.024144	.093698	.011553
Social	.0254768	.03373	0.76	0.450	-.040632	.091585	-.052786
Career	-.0635838	.04101	-1.55	0.121	-.143969	.016801	.048372
Cope	-.0015287	.03929	-0.04	0.969	-.078545	.075488	-.07255

(*) dy/dx is for discrete change of dummy variable from 0 to 1

Table 49 Multinomial logit analysis of education path with respect to four factors, Female dummy variable and cognitive abilities, family and background variables

```
Iteration 0:   log pseudolikelihood = -404.83941
Iteration 1:   log pseudolikelihood = -305.19615
Iteration 2:   log pseudolikelihood = -287.70249
Iteration 3:   log pseudolikelihood = -285.78721
Iteration 4:   log pseudolikelihood = -285.54829
Iteration 5:   log pseudolikelihood = -285.49414
Iteration 6:   log pseudolikelihood = -285.47615
Iteration 7:   log pseudolikelihood = -285.27226
Iteration 8:   log pseudolikelihood = -285.24252
Iteration 9:   log pseudolikelihood = -285.24218
Iteration 10:  log pseudolikelihood = -285.24188
Iteration 11:  log pseudolikelihood = -285.24188

Multinomial logistic regression               Number of obs     =        215
                                              Wald chi2(83)     =          .
                                              Prob > chi2       =          .
Log pseudolikelihood = -285.24188             Pseudo R2         =     0.2954
```

jobfield	RRR	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
Physical_math_engineering_prof_a	(base outcome)					
Teaching_prof__and_associates						
medu	4.800538	7.953117	0.95	0.344	.1866779	123.4488
fedu	27.60176	35.65178	2.57	0.010	2.195248	347.0485
urban	.9522287	.9981995	-0.05	0.963	.1220243	7.430812
wlemath	.9871133	.0095552	-1.34	0.180	.968562	1.00602
wleread	1.010855	.0077302	1.41	0.158	.9958168	1.02612
Age	.9867913	.4057705	-0.03	0.974	.4407675	2.20923
female	4.320116	6.487588	0.97	0.330	.2276314	81.98957
wealth	.3355802	.2540478	-1.44	0.149	.0761031	1.479757
Esteem	.5123533	.2462637	-1.39	0.164	.1997274	1.314321
Social	4.992416	2.96983	2.70	0.007	1.555815	16.02004
Career	1.074939	.6594424	0.12	0.906	.3229963	3.577421
Cope	2.419355	1.621762	1.32	0.187	.6503024	9.00086
_cons	.2532006	2.289689	-0.15	0.879	5.08e-09	1.26e+07
Social_sciences_Business_law_pro						
medu	1.798956	2.852318	0.37	0.711	.0804278	40.23789
fedu	2.911654	3.602895	0.86	0.388	.2575483	32.91704
urban	2.15243	1.400814	1.18	0.239	.6011212	7.70719
wlemath	.9943446	.0050508	-1.12	0.264	.9844943	1.004294
wleread	.9981271	.0063426	-0.30	0.768	.9857729	1.010636
Age	.9866618	.405717	-0.03	0.974	.4407099	2.20894
female	2.482255	2.050946	1.10	0.271	.4915267	12.53562
wealth	.733632	.3401422	-0.67	0.504	.2956827	1.820248
Esteem	.80375	.3303814	-0.53	0.595	.3591159	1.798902
Social	1.468023	.7934056	0.71	0.477	.5089717	4.234206
Career	1.573645	.9138521	0.78	0.435	.5041904	4.911557
Cope	1.546434	.917404	0.73	0.462	.483466	4.946486
_cons	51.84163	377.4182	0.54	0.588	.0000329	8.16e+07
Life_and_health_science_prof__an						
medu	1.643592	3.20624	0.25	0.799	.0359179	75.21019
fedu	3.933667	5.569024	0.97	0.333	.2453145	63.07715
urban	.2869457	.2382683	-1.50	0.133	.0563632	1.460843
wlemath	.995406	.0066863	-0.69	0.493	.982387	1.008598
wleread	.9979794	.0070647	-0.29	0.775	.9842283	1.011923
Age	.9867871	.4057922	-0.03	0.974	.4407451	2.209324
female	2.25e+08	1.89e+08	22.93	0.000	4.34e+07	1.16e+09
wealth	1.014896	.8499125	0.02	0.986	.1966043	5.239021
Esteem	1.033535	.4981403	0.07	0.945	.4018512	2.658183
Social	6.027438	4.134513	2.62	0.009	1.571275	23.12136
Career	2.6	2.019814	1.23	0.219	.5671741	11.91874
Cope	1.923076	1.177089	1.07	0.285	.5794122	6.382714
_cons	3.47e-07	2.77e-06	-1.86	0.063	5.46e-14	2.206418

clerks_office_clerks						
medu	.7656578	1.189595	-0.17	0.864	.0364361	16.08933
fedu	6.172985	7.168222	1.57	0.117	.6339558	60.10788
urban	1.978491	1.149171	1.17	0.240	.6337654	6.176456
wlemath	.9934549	.004347	-1.50	0.133	.9849714	1.002011
wleread	.9939312	.0056745	-1.07	0.286	.9828714	1.005115
Age	.9857573	.405364	-0.03	0.972	.4402893	2.206997
female	3.484233	2.397619	1.81	0.070	.9044122	13.42295
wealth	.7924746	.357355	-0.52	0.606	.3274525	1.917884
Esteem	.7607744	.2612441	-0.80	0.426	.3881145	1.491255
Social	1.337073	.5603277	0.69	0.488	.5880873	3.039964
Career	3.076089	1.211956	2.85	0.004	1.421124	6.658341
Cope	2.364284	1.300676	1.56	0.118	.8043105	6.949851
_cons	1743.169	12377.63	1.05	0.293	.0015749	1.93e+09
service_workers						
medu	2.272551	3.838488	0.49	0.627	.0829427	62.26573
fedu	2.656628	3.573107	0.73	0.468	.1903195	37.08329
urban	.7752612	.5490597	-0.36	0.719	.1934664	3.106638
wlemath	.9905444	.0054922	-1.71	0.087	.9798382	1.001368
wleread	.9807187	.0062105	-3.07	0.002	.9686216	.9929669
Age	1.282075	.6406314	0.50	0.619	.4814841	3.413851
female	4.055378	3.969197	1.43	0.153	.5955509	27.61493
wealth	.7025963	.3679178	-0.67	0.500	.2517504	1.960837
Esteem	1.405462	.5993843	0.80	0.425	.609269	3.242121
Social	3.576518	1.955327	2.33	0.020	1.224892	10.44294
Career	2.558907	1.234465	1.95	0.051	.9940747	6.587037
Cope	2.50892	1.501685	1.54	0.124	.7762647	8.108934
_cons	36691.93	317267.2	1.22	0.224	.0016011	8.41e+11
agriculture						
medu	3.12e-08	4.50e-08	-12.00	0.000	1.86e-09	5.25e-07
fedu	2.50e-08	3.23e-08	-13.58	0.000	2.00e-09	3.13e-07
urban	1.615271	1.438741	0.54	0.590	.2818838	9.25594
wlemath	.9878213	.0056045	-2.16	0.031	.9768976	.9988672
wleread	.9928237	.0087365	-0.82	0.413	.9758472	1.010096
Age	.9863027	.4056014	-0.03	0.973	.4405214	2.208276
female	2.107843	1.955893	0.80	0.422	.341978	12.99207
wealth	.1570641	.1232813	-2.36	0.018	.0337256	.7314657
Esteem	1.448006	.5945296	0.90	0.367	.6475598	3.237882
Social	.6313244	.3310018	-0.88	0.380	.2259278	1.76415
Career	1.351542	.7510151	0.54	0.588	.4548207	4.016234
Cope	.589019	.379294	-0.82	0.411	.1667267	2.080911
_cons	4955.938	40715.91	1.04	0.300	.0005035	4.88e+10
craft__plant__and_machine_operat						
medu	3.306993	4.965738	0.80	0.426	.1742931	62.74606
fedu	3.645376	4.27509	1.10	0.270	.3660172	36.30639
urban	.6349701	.4048931	-0.71	0.476	.1819612	2.215786
wlemath	1.003024	.0044853	0.68	0.500	.9942711	1.011853
wleread	.9806752	.0059161	-3.23	0.001	.9691482	.9923393
Age	1.083768	.5590286	0.16	0.876	.394339	2.978534
female	.151911	.120678	-2.37	0.018	.0320181	.7207474
wealth	.5368356	.2742808	-1.22	0.223	.197217	1.461297
Esteem	1.191852	.4431171	0.47	0.637	.5751161	2.469958
Social	1.897326	.8107241	1.50	0.134	.8211546	4.383882
Career	1.282493	.5802216	0.55	0.582	.5283926	3.112813
Cope	2.191275	1.230201	1.40	0.162	.729163	6.585201
_cons	5915.101	54190.09	0.95	0.343	.0000942	3.72e+11

Note: _cons estimates baseline relative risk for each outcome.

Table 50 M.E. on probability of choosing physical, math and engineering professions

```
. mfx, predict(pr outcome(1))

Marginal effects after mlogit
      y  = Pr(jobfield==Physical_math_engineering_prof_a) (predict, pr outcome(1))
         =  .06783135
```

variable	dy/dx	Std. Err.	z	P>|z|	[95%	C.I.]	X
medu*	-.0236388	.06732	-0.35	0.726	-.155592	.108315	.12093
fedu*	-.0711597	.03567	-1.99	0.046	-.141072	-.001248	.218605
urban*	-.0231221	.03551	-0.65	0.515	-.092729	.046485	.553488
wlemath	.0003598	.00026	1.37	0.172	-.000156	.000876	540.482
wleread	.0005508	.00036	1.53	0.125	-.000154	.001255	518.761
female*	-.0336111	.03712	-0.91	0.365	-.106374	.039152	.590698
wealth	.0216166	.02784	0.78	0.438	-.032956	.076189	-.037581
Esteem	.0077019	.01974	0.39	0.696	-.030991	.046395	.011553
Social	-.0333675	.02879	-1.16	0.247	-.0898	.023065	-.052786
Career	-.053539	.02418	-2.21	0.027	-.100936	-.006141	.048372
Cope	-.0498385	.03277	-1.52	0.128	-.114062	.014385	-.07255

(*) dy/dx is for discrete change of dummy variable from 0 to 1

Table 51 M.E. on probability of choosing teaching professions

```
Marginal effects after mlogit
      y  = Pr(jobfield==Teaching_prof__and_associates) (predict, pr outcome(2))
         =  .01687989
```

variable	dy/dx	Std. Err.	z	P>|z|	[95%	C.I.]	X
medu*	.0324194	.03866	0.84	0.402	-.043343	.108182	.12093
fedu*	.0575708	.03579	1.61	0.108	-.012581	.127723	.218605
urban*	-.0065584	.01775	-0.37	0.712	-.041353	.028236	.553488
wlemath	-.0001321	.00019	-0.68	0.495	-.000511	.000247	540.482
wleread	.0003199	.00023	1.37	0.171	-.000138	.000777	518.761
female*	.011804	.01553	0.76	0.447	-.018626	.042234	.590698
wealth	-.0129289	.01106	-1.17	0.243	-.034613	.008755	-.037581
Esteem	-.0092955	.00743	-1.25	0.211	-.023862	.005271	.011553
Social	.0184176	.01411	1.31	0.192	-.009233	.046068	-.052786
Career	-.0126563	.01277	-0.99	0.322	-.037688	.012375	.048372
Cope	.0027166	.00778	0.35	0.727	-.012536	.017969	-.07255

(*) dy/dx is for discrete change of dummy variable from 0 to 1

Table 52 M.E. on probability of choosing social science, business, law professions

```
Marginal effects after mlogit
      y  = Pr(jobfield==Social_sciences_Business_law_pro) (predict, pr outcome(3))
         =   .1692021
```

variable	dy/dx	Std. Err.	z	P>|z|	[95%	C.I.]	X
medu*	.0333302	.104	0.32	0.749	-.170502	.237162	.12093
fedu*	-.0640269	.07498	-0.85	0.393	-.210981	.082927	.218605
urban*	.0706066	.06528	1.08	0.279	-.057339	.198552	.553488
wlemath	-.0000647	.00052	-0.12	0.901	-.00108	.00095	540.482
wleread	.0010514	.0005	2.10	0.036	.000071	.002032	518.761
female*	.0479719	.07432	0.65	0.519	-.097696	.19364	.590698
wealth	.0026301	.0357	0.07	0.941	-.067348	.072608	-.037581
Esteem	-.0174927	.04566	-0.38	0.702	-.106992	.072007	.011553
Social	-.0202424	.05146	-0.39	0.694	-.121112	.080627	-.052786
Career	-.0580408	.06287	-0.92	0.356	-.181268	.065186	.048372
Cope	-.0494375	.04173	-1.18	0.236	-.131227	.032352	-.07255

(*) dy/dx is for discrete change of dummy variable from 0 to 1

Table 53 . M.E. on probability of choosing health science professions

```
. mfx, predict(pr outcome(4))

Marginal effects after mlogit
      y  = Pr(jobfield==Life_and_health_science_prof__an) (predict, pr outcome(4))
         =  .00012379
```

variable	dy/dx	Std. Err.	z	P>|z|	[95%	C.I.]	X
medu*	7.85e-06	.00019	0.04	0.966	-.000358	.000374	.12093
fedu*	-.0000148	.0001	-0.14	0.887	-.000219	.00019	.218605
urban*	-.0002195	.00012	-1.82	0.068	-.000456	.000017	.553488
wlemath	8.98e-08	.00000	0.13	0.894	-1.2e-06	1.4e-06	540.482
wleread	7.02e-07	.00000	1.22	0.223	-4.3e-07	1.8e-06	518.761
female*	.053244	.02869	1.86	0.063	-.002983	.109471	.590698
wealth	.000045	.00008	0.54	0.588	-.000118	.000208	-.037581
Esteem	.0000225	.00005	0.47	0.639	-.000072	.000117	.011553
Social	.000153	.00008	2.04	0.042	5.9e-06	.0003	-.052786
Career	3.80e-06	.00008	0.05	0.961	-.000148	.000156	.048372
Cope	-9.13e-06	.00004	-0.20	0.838	-.000097	.000078	-.07255

Table 54 M.E. on probability of choosing offcieclerks professions

```
. mfx, predict(pr outcome(5))

Marginal effects after mlogit
      y  = Pr(jobfield==clerks_office_clerks) (predict, pr outcome(5))
         =  .48741128
------------------------------------------------------------------------------
variable |      dy/dx    Std. Err.     z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
    medu*|  -.2461878       .1297   -1.90   0.058   -.50039  .008014    .12093
    fedu*|   .1550098      .10934    1.42   0.156  -.059284  .369303   .218605
   urban*|   .1532677      .08816    1.74   0.082  -.019527  .326062   .553488
 wlemath |  -.0006651       .0007   -0.95   0.343  -.002041  .000711   540.482
 wleread |   .0011055      .00069    1.60   0.109  -.000245  .002456   518.761
  female*|   .2634225      .08834    2.98   0.003   .090283  .436562   .590698
  wealth |   .0371799      .05281    0.70   0.481  -.066333  .140693  -.037581
  Esteem |  -.0787603      .05546   -1.42   0.156  -.187465  .029945   .011553
  Social |  -.0924411      .05792   -1.60   0.110  -.205957  .021075  -.052786
  Career |   .1816073      .06857    2.65   0.008   .047212  .316003   .048372
    Cope |    .067266      .05858    1.15   0.251  -.047554  .182086   -.07255
------------------------------------------------------------------------------
(*) dy/dx is for discrete change of dummy variable from 0 to 1
```

Table 55 M.E. on probability of service workers professions

```
Marginal effects after mlogit
      y  = Pr(jobfield==service_workers) (predict, pr outcome(6))
         =  .12929897
------------------------------------------------------------------------------
variable |      dy/dx    Std. Err.     z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
    medu*|   .0642581      .13974    0.46   0.646  -.209621  .338137    .12093
    fedu*|  -.0513634      .06728   -0.76   0.445  -.183231  .080505   .218605
   urban*|   -.083985      .06001   -1.40   0.162  -.201604  .033634   .553488
 wlemath |  -.0005716      .00047   -1.22   0.222  -.001489  .000346   540.482
 wleread |  -.0015374      .00045   -3.43   0.001  -.002415  -.00066   518.761
  female*|    .080065      .06034    1.33   0.185  -.038196  .198326   .590698
  wealth |  -.0074529      .03407   -0.22   0.827  -.074233  .059328  -.037581
  Esteem |    .059968      .03897    1.54   0.124  -.016409  .136344   .011553
  Social |    .106151      .04739    2.24   0.025   .013264  .199037  -.052786
  Career |   .0136091       .0381    0.36   0.721  -.061057  .088275   .048372
    Cope |   .0231993       .0364    0.64   0.524  -.048137  .094536   -.07255
------------------------------------------------------------------------------
(*) dy/dx is for discrete change of dummy variable from 0 to 1
```

Table 56 M.E. on probability of choosing agriculture professions

```
Marginal effects after mlogit
      y  = Pr(jobfield==agriculture) (predict, pr outcome(7))
         =  .00016817
```

variable	dy/dx	Std. Err.	z	P>|z|	[95% C.I.	]	X
medu*	-.0009342	.00067	-1.39	0.164	-.002251	.000383	.12093
fedu*	-.0049589	.00329	-1.51	0.132	-.011413	.001495	.218605
wealth	-.0002563	.00012	-2.09	0.037	-.000497	-.000016	-.037581
urban*	.0000226	.00012	0.18	0.856	-.000221	.000266	.553488
wleread	1.43e-07	.00000	0.12	0.907	-2.2e-06	2.5e-06	518.761
wlemath	-1.17e-06	.00000	-1.42	0.154	-2.8e-06	4.4e-07	540.482
female*	.000026	.00011	0.25	0.806	-.000182	.000234	.590698
Esteem	.0000836	.00006	1.40	0.161	-.000033	.0002	.011553
Social	-.0001601	.00009	-1.74	0.081	-.00034	.00002	-.052786
Career	-.0000807	.0001	-0.77	0.440	-.000286	.000124	.048372
Cope	-.0002106	.00011	-1.96	0.050	-.000421	2.5e-07	-.07255

(*) dy/dx is for discrete change of dummy variable from 0 to 1

Table 57 M.E. on probability of choosing craft, plant and machine operating professions

```
Marginal effects after mlogit
      y  = Pr(jobfield==craft__plant__and_machine_operat) (predict, pr outcome(8))
         =  .12908447
```

variable	dy/dx	Std. Err.	z	P>|z|	[95% C.I.	]	X
medu*	.1407453	.15131	0.93	0.352	-.155822	.437313	.12093
fedu*	-.0210568	.05779	-0.36	0.716	-.134329	.092215	.218605
wealth	-.0408336	.03591	-1.14	0.255	-.111208	.029541	-.037581
urban*	-.1100119	.06192	-1.78	0.076	-.231378	.011354	.553488
wleread	-.001491	.0004	-3.73	0.000	-.002275	-.000707	518.761
wlemath	.0010748	.00037	2.89	0.004	.000347	.001803	540.482
female*	-.4229223	.08077	-5.24	0.000	-.581219	-.264626	.590698
Esteem	.0377727	.03055	1.24	0.216	-.022096	.097642	.011553
Social	.0214896	.03469	0.62	0.536	-.046508	.089487	-.052786
Career	-.0709034	.04135	-1.71	0.086	-.151949	.010142	.048372
Cope	.0063138	.03681	0.17	0.864	-.065829	.078457	-.07255

(*) dy/dx is for discrete change of dummy variable from 0 to 1