Johann Wolfgang Goethe Universität Frankfurt am Main

Winter Semester 2014/2015

Student Advisor: Prof. Dr. Cornelia Storz,

Chair for the Study of Economic Institutions, Innovation and East Asian Development

2nd Reviewer: Dr. Peter Ganea, Program Coordinator MEAS

Topic:

Job Satisfaction - An Econometric Analysis based on the Japanese General Social Survey 2010

Thesis
submitted in partial fulfillment of the requirements for the degree
Master of Arts in Modern East Asian Studies

by:
Katharina Ditte

Contents

List of Figures

List of Tables

List of Symbols

ESS explained sum of squares. 108

$McFadden - Pseudo - R^2$ likelihood-ratio index (R^2 for logit models). 108

RSS residual sum of squares. 108

R^2 coefficient of determination. 107

TSS total sum of squares. 108

X_i vector of the explanatory variables for individual i. 27

$\bar{R}^2$ coefficient of determination adjusted for degrees of freedom. 35

β parameter vector. 27

λ_k^i threshold of category k for individual i. 27

$\sum\left(Y_i - \bar{Y}\right)$ total sum of squares. 108

$\sum e_i^2$ residual sum of squares. 108

ε_i error term for individual i. 27

i individual i. 27

js self-reported job satisfaction. 27

js^* latent job satisfaction. 27

k number of category. 27

$lnL(\hat{\theta})$ log likelihood of the full model. 108

$lnL(\tilde{\theta})$ log likelihood of the intercept model. 108

x explanatory variable x. 28

y dependent variable y. 28

z explanatory variable z (distinct from variable x). 33

max. maximum. 71, 75, 79, 84

min. minimum. 71, 75, 79, 84

n number of observations. 71, 75, 79, 84

qu. quartile. 71, 75, 79, 84

List of Abbreviations

2SLS two-stage least squares. 29

ag. agency. 105

com. commuting. 105

df degrees of freedom. 41, 91

disp. dispatched. 104

EFILWC European Foundation for the Improvement of Living and Working Conditions. 10

emp. Employed. 105

EU European Union. 3

g. general. 104

GDP gross domestic product. 5

gov. government. 105

IV instrumental-variable. 29

JGSS Japanese General Social Survey. 3

m. men. 104

MAR missing at random. 34

no. number of children. 104

OECD Organization for Economic Co-operation and Development. 7

OLS ordinary least squares. 3

org. organization. 105

sat. satisfaction. 105

SEs standard errors. 32

SRG stress-related growth. 15

temp. temporary. 104

US United States. 3

USA United States of America. 5

VIF variance inflation factor. 33

w. women. 104

1 Introduction

1.1 Background and Need

For centuries happiness has been mainly a subject of poetry and philosophical works, but in the last few decades economists have raised their interest in the topic. As a result the new academic subfield - happiness economics - has emerged. Researchers of this field use statistical methods to find, validate and analyze the relationship between different measures of happiness (also described as subjective well-being) and demographic, economic and psychological variables.

In contrast to poets and philosophers, economists usually don't go into further detail on the meaning of happiness. Instead they use it as a proxy measure of utility (e.g. Frey & Stutzer, 2002a; Pugno, 2007, p. 263; Royo, 2007, p. 160) assuming that people are able to provide a more or less accurate self-assessment of their happiness level on a scale of ordinal values (e.g. from 1 to 10). This approach is legitimized by a large body of psychology literature providing evidence on the reliability of such data through the application of a variety of validation tests. These tests show that self-assessments of happiness strongly correlate with objective measures of well-being, such as the frequency of smiling, neural activity, heart rate or hormone levels (e.g. Ekman, Davidson, & Friesen, 1990; Pavot & Diener, 1993; Sandvik, Diener, & Seidlitz, 1993 ; Sutton & Davidson, 1997). Particularly, Veenhoven has responded in detail to concerns over the scientific validity of self-reported happiness data (Veenhoven, 1994; Veenhoven, 1996; Veenhoven, 1997) with the conclusion that most objections "can be discarded" (Veenhoven, 1996, p. 3).

In labour economics, the rise of this new academic field happiness economics has inspired a growing number of studies on job satisfaction (D'Addio, Eriksson, & Frijters, 2007, p. 2413). Similar to the common economic definition of happiness,[1] job satisfaction can be described as the overall degree of satisfaction with one's

[1] In most economic studies *"happiness"* is broadly defined as the degree of overall satisfaction with the quality of one's life (Veenhoven, 1991, p. 2; Blanchflower & Oswald, 2004, p. 1360; Graham & Felton, 2005, p. 108).

job. In labour economics it is used as a proxy measure of utility received from work (e.g. Clark, 1997, p. 344; Benz & Frey, 2008, p. 362; Graaf-Zijl, 2011, p. 198; Poggi, 2014, p. 1). It is usually derived from surveys, which asks their participants to estimate their level of satisfaction on an integer scale (e.g. from 1 to 5) or by choosing a descriptive category (e.g. very dissatisfied, fairly dissatisfied, neither nor, fairly satisfied, very satisfied) (Sousa-Poza & Sousa-Poza, 2000, p. 523; Clark, 2003; Ball & Chernova, 2007, p. 498).

Besides being an important predictor of life satisfaction (e.g. Judge & Watanabe, 1993; Clark, 2010; Bowling, Eschleman, & Wang, 2010), job satisfaction plays an meaningful role in the performance of businesses (EFILWC, 2007, p. 2; Nielsen & Smyth, 2008, p. 1921). For instance, high job satisfaction can help to prevent burnout (Tourigny, Baba, & Wang, 2010) and decrease labour turnover and absenteeism among employees (Clark, 2001; Mobley, 1977; Lee, Trevor, Gerhart, & Weller, 2008; Jones, Jones, Latreille, & Sloane, 2009, p. 5). Moreover, it is associated with high organizational commitment (Mathieu & Zajac, 1990, p. 183; Srivastava, 2013). In addition to these economic reasons managers and employers often have a moral motivation to contribute to higher job satisfaction in their company. Being aware of the groups who are especially at risk of suffering from low job satisfaction and knowing potential means to increase work satisfaction (e.g. pay raise, team building events, leisure programs, etc.) can help them to reach this goal more effectively - irrespective of whether their primary reasons are economic or normative.

However, the body of literature on the relation between job satisfaction and classic demographic and economic determinants - such as age, gender, marital status and so forth - is still relatively small compared to the amount of material existing on general life satisfaction. In fact, some very common determinants in happiness research have rarely or never been studied in context with job satisfaction.

One of those missing fields is the role of leisure activities and leisure satisfaction. However, more research on this topic is desirable, because in contrast to fixed factors like gender, age or personality, it is easily possible for people to influence and change their level of leisure satisfaction through their choice of activities (Argyle, 2013, p. 112). Therefore, more research in this field would be beneficial for institutions (e.g. companies, governmental organizations, etc.) as well as individuals. More specifically, new findings in this area could help employees to pick the right activities in order to cope more effectively with stressful or unpleasant parts of their job, while companies could receive valuable information to design better leisure programs to improve the work life balance and job satisfaction of their staff.

The lack of literature on the determinants of job satisfaction is especially apparent in respect to regions outside the European Union (EU) and the United States (US). There is a special need to close this remaining gap, because findings on different determinants of job satisfaction frequently differ depending on the characteristics of the sample[2] or the time the survey was conducted.[3] This means that we cannot simply assume that findings from one cultural setting (e.g. an Anglo-Saxon or European country) will also apply to a different cultural context (e.g. an African or Asian country).

1.2 Purpose of the Study and Applied Method

With this thesis the researcher aims to contribute to the closure of the above stated research gaps by providing a detailed overview of the influence of different social and subjective well-being variables on job satisfaction in Japan. In particular the study attempts to shed light on the the link between job satisfaction and leisure satisfaction as well as different forms of recreation.

To achieve this goal a regression analysis assessing the relationship between job satisfaction and a set of different social and subjective well-being variables as well as different types of leisure activities is conducted. The representative nationwide

[2]E.g. results can heavily differ depending on the cultural context, type of employment or gender of the survey participants.

[3]E.g. older studies tend to find a positive linear relationships between age and job satisfaction, while studies conducted after 2000 more frequently find a U-shaped relationship.

Japanese General Social Survey (JGSS) is used as the data source. Because the survey measures overall job satisfaction on a 5-point ordinal scale, an ordered logit model is estimated - an approach which is commonly applied by economists on data with the same characteristics. However, since psychological studies tend to use ordinary least squares (OLS) regression instead, an additional OLS model is estimated. This helps to prove the persistence of the observed correlation and makes the findings more comparable to previous and future research. Furthermore, since a number of previous studies (e.g. Sousa-Poza & Sousa-Poza, 2000) on job satisfaction have observed differences in the estimates between women and men, additional regressions are run for the male and female sample.

2 Review of the Literature

2.1 Background

The economic literature on happiness started with a study by Easterlin (1974), which documented stagnant average happiness levels in the United States of America (USA), despite large increases of the gross domestic product (GDP) during the same period. This phenomenon - also known as the "Easterlin Paradox" - still receives frequent attention by recent researchers (e.g. Inglehart, 1996; Blanchflower & Oswald, 2004; Vendrik & Hirata, 2007; Clark, Frijters, & Shields, 2008). Since this seminal paper, the body of empirical economic research using self-reported happiness data has risen significantly. In addition, the birth of *"happiness economics"* had spill-over effects on other subfields of economics. One example is labour economics, were the body of literature on job satisfaction started to grow shortly after Easterlin published his famous article.[4]

2.2 Determinants of Job Satisfaction

Before the models of this study are introduced, it is important to explain the major determinants, which come into questions as controls or independent variables. This section introduces the major determinants of job satisfaction which have been identified by previous studies. Compared with the results on general life satisfaction or happiness for which the reported effects are mostly consistent - even over time and across different countries[5] - there are more conflicting findings on job satisfaction. This observation indicates that job satisfaction might strongly depend on the institutional or cultural setting.

[4]Many economics articles on job satisfaction cite Hammermesh (1977) and Freeman (1978) as the seminal works.

[5]For additional information an introduction to the field of happiness economics in book-length is provided by Frey and Stutzer (2002b).

2.2.1 Social Variables

Social variables describe characteristics of survey participants which may impact the way they respond to certain questions. Generally, social variables capture demographic information (e.g. age, gender, legal marital status etc.), socio-economic information (e.g. type of occupation, monthly income, education level etc.) and, geographic information (e.g. country of residence, degree of urbanization etc.) (European Commission, 2007). In respect to job satisfaction previous research has identified the following relationships:

Age: Several studies find a positive relation between age and job satisfaction in the USA (Weaver, 1978, p. 272; Near, Rice, & Hunt, 1978, p. 255; Loscocco, 1990, p. 165), China (Loscocco & Bose, 1998, p. 102) and Russia (Linz, 2004, p. 275). Furthermore, Rhodes (1983) finds a linear positive correlation between job satisfaction and age for the vast majority of studies (79%) conducted before 1983. The so-called *"grinding down hypothesis"* is one of the most common explanations for a linear increase in job satisfaction with age (Wright & Hamilton, 1978, p. 1142): It assumes that young workers enter the labour market with too high expectations. The fact that their first job often does not comply with their hopes causes dissatisfaction. However, after years of experience they adjust their aspirations and grow happier with their job. In a nutshell, higher job satisfaction over time is mainly caused by lower expectations rather than an factual improvement in working conditions (Fasang, Geerdes, Schönmann, & Siarov, 2007, p. 5).
However, several recent studies indicate a U-shaped relationship in Great Britain (Clark, Oswald, & Warr, 1996), Europe (Bauer, 2004, p. 12), the USA (Bender, 2005, p. 486), Kazakhstan, Russia and Armenia (Linz & Semykina, 2012, p. 829). An explanation for the U-shaped relationship is that employees are often euphoric at the beginning of their career due to new experiences at their job and the feeling of finally being independent. Yet, this energizing effect diminishes gradually and so does job satisfaction (Fasang et al., 2007, p. 6). If their assignments get more

interesting as their career progresses and they start to gain more responsibilities, this effect is eventually revered and job satisfaction starts rising again at a certain age. The *"developmental aging hypothesis"* provides an alternative explanation for the U-shaped relationship. It builds on the assumption of alternating stable and transitional phases in life. During a transition phase people may look back, question and adjust their previous life views. It is expected that especially middle-aged people will struggle during the reevaluating process, and as a result their job satisfaction might drop. Older people - being more mature and experienced in personality development - are presumed to cope better with a transition phase (Fasang et al., 2007, p. 6). Another possibility is that it is not job satisfaction, but an omitted variable that causes the U-shaped pattern. A promising candidate is general *self-reported happiness*, which is expected to be positively related to job satisfaction and was reported in a vast majority of works from the field of happiness studies to have a U-shaped relationship with age (Horley & Lavery, 1995; Frey & Stutzer, 2002b, p. 53).

Gender: There is evidence that women in the USA (Hodson, 1989; Loscocco, 1990, p. 162; Bender, 2005), Great Britain (Clark et al., 1996; Clark, 1997; Asadullah & Fernández, 2008), Japan (Lincoln, 1992, p. 159) and other member countries of the Organization for Economic Co-operation and Development (OECD) (Clark, 1998, p. 16) tend to have higher job satisfaction scores or at least about the same as men (Weaver, 1978, p. 272*f*, USA; Mannheim, 1983, p. 423, Israel), despite being clearly disadvantaged in terms of earnings, recruitment, and promotion (e.g. Clark, 1997, p. 342; EFILWC, 2007, p. 12). There are several different explanations for this *"paradox of the contented female worker"* (Bender, 2005, p. 479; Asadullah & Fernández, 2008, p. 1):

1. Lower expectations: Some researchers assume that female employees just have lower expectation (e.g. Ferree, 1976; Lincoln, 1992, p. 159) than their male counterparts, which might result from their historically weak position in the labour market (Clark, 1997, p. 342). This view is supported by interaction experiments showing that the gender satisfaction differential does not hold in the case of young or highly-educated employees as well as for those working in male-dominated occupations - all groups are characterized by relatively low gender differences in job expectations (Clark, 1997, p. 342). Therefore it is expected that the gap will decrease over time, when women adjust their expectations to the improvements. Indeed, the gender gap seems to have decreased in recent years in Great Britain (Asadullah & Fernández, 2008, p. 13).

2. Different reference group: One other explanation is that women tend to compare themselves to other women, while men compare themselves to other men. And since, there are less women in leading and high paying positions, this might lead to lower aspiration levels among women (Crosby, 1982; Varce, Shaffer, & Cynthia, 1983, p. 352). One fact supporting this view is that women tend to be more satisfied with their job, if their mothers stayed at home and if they are working in workplaces with a majority of female employees (Hodson, 1989, p. 397). However, the last finding was challenged by Bender (2005), who demonstrated that if differences in job flexibility are included as controls, the relationship between the gender composition of the workplace and job satisfaction of women disappears.

3. Self selection: There are at least two versions of this argument. In the first version it is assumed that women tend to self select themselves into jobs which provide a greater amount of work-life balance. Indeed, in a study by Asadullah & Fernández (Asadullah & Fernández, 2008) the gender gap in job satisfaction decreased by around a half, after work-life balance was included as a control.

In the second version it is argued that women who do not find fulfillment at work are more likely to self select themselves into housework, while women with a high job satisfaction tend to remain in the paid labour market (Clark, 1997, p. 345). However, a number of studies do not support the “paradox of the contented female worker” and show an opposite pattern - higher job satisfaction levels among men

- for China (Loscocco & Bose, 1998), Japan (Boyles & Shibata, 2009, p. 59), Portugal and Denmark (EFILWC, 2007, p. 12). Furthermore, some studies do not find a significant relationship between gender and job satisfaction in the USA (e.g. Near et al., 1978, p. 255).

Marital status: Married employees tend to have a higher job satisfaction in Japan (Lincoln, 1992, p. 159), Great Britain (Clark et al., 1996, p. 65), Denmark, Italy, the Netherlands (EFILWC, 2007, p. 13) and the USA (Weaver, 1978, p. 272; Loscocco, 1990, p. 166). One possible explanation for this could be that married individuals don't depend only on their own income, but can rely on the support of their partner in difficult times. This allows them to select a job they truly like, while individuals who are single are more likely to be forced to take jobs which provide only low satisfaction (Nielsen & Smyth, 2008, p. 1928). One other explanation is that individuals in a committed relationship may have more resources to draw on (i.e., their spouse, more finances) contributing to lower levels of stress (Grandey & Cropanzano, 1999, p. 353).
However, being in a committed relationship might also reduce one's freedom to choose from different job opportunities, for example by restricting one's choice only to a certain area or city. So the coefficient of the variable depends on the relative importance of positive and negative effects. Indeed, one study shows that in Austria, Bulgaria, Germany and Portugal the *"never married"* group exhibits the highest levels of job satisfaction and in Romania the *"single/unmarried"* group was the one with the highest job satisfaction levels (EFILWC, 2007, p. 13). There are also studies finding the highest job satisfaction rates among the *"widowed"* group in the USA (Near et al., 1978, p. 257).

Number of children: Previous studies have shown that the number of children is significantly and negatively related to work-family conflicts (Eagle, Miles, & Icenogle, 1997, p. 169; Grandey & Cropanzano, 1999, p. 353). This led to the

assumption that the number of dependent children could negatively influence job satisfaction (Nielsen & Smyth, 2008, p. 1927).
However, most studies could not find any empirical support for this hypothesis. In fact, the majority of studies find no significant effect of having children and level of job satisfaction. For instance, no effect of young or dependent children on job satisfaction was found in the USA (Hanson & Sloane, 1992) and China (Nielsen & Smyth, 2008, p. 1998) and no significant effect of being a parent and job satisfaction was observed in Czech Republic, Denmark and Italy (EFILWC, 2007, p. 13). Yet, Clark et al. (1996, p. 65) get an ambiguous result for Great Britain: Although a negative insignificant effect of having one or two children and job satisfaction was measured, the same study found also a positive and significant effect of having three and more children. Finally, some data from the Netherlands, Romania and Finland even suggests that job satisfaction of parents is higher and even increases with the number of children (EFILWC, 2007, p. 13).

Education: Previous studies have found a negative effect of higher education on job satisfaction (when controlled for income) in the USA (Hodson, 1989, p. 393; Loscocco, 1990, p. 166; Bender, 2005, p. 486), Great Britain (Clark et al., 1996; Asadullah & Fernández, 2008, p. 12), Japan (Lincoln, 1992, p. 157) and China (Loscocco & Bose, 1998, p. 102). One explanation for this observation is that education raises expectations considering intrinsic (e.g. challenging and interesting assignments) and extrinsic (e.g. income and fringe benefits) job rewards, which causes low job satisfaction, if these expectations do not resonate with the reality (Loscocco & Bose, 1998, p. 102; Nielsen & Smyth, 2008, p. 1926).
However, Bender (2005, p. 486) finds that the high significance largely depends on the male subsample (for women higher education has a negative, but not significant effect).
But there are also contradicting findings: For instance a study by the European Foundation for the Improvement of Living and Working Conditions (EFILWC)

(2007, p. 10) finds a positive effect of education on job satisfaction for most European countries. Positive effects of education are likely, if the institutional setting favors highly educated employees in terms of occupational mobility, risk of unemployment and promotional chances (Fasang et al., 2007, p. 8).

Respondent's income: It is not surprising that a positive relationship between wage and job satisfaction is supported by a vast body of empirical research (e.g. for the USA: Weaver, 1974, p. 374, Hodson, 1989, p. 392; for Great Britain: Asadullah & Fernández, 2008, p. 25; for Denmark: D'Addio et al., 2007, p. 2420; for Armenia, Kazakhstan, Kyrgyzstan, Russia, and Serbia: Linz & Semykina, 2012, p. 836; for Europe: Bauer, 2004, p. 12 and for China: Loscocco & Bose, 1998, p. 106). In addition, Pouliakas & Theodossiou (2010) find that high-paid employees are significantly more satisfied with their jobs than low-paid employees in Greece, Ireland, Italy, Portugal and Spain.

In addition, Clark (1998, p. 12), Asadullah & Fernández (2008, p. 25) and D'Aggio et al. (2007) show that the significance of the wage variable is stronger for men than for women. This result is consistent with the common finding in the psychology and management literature that women tend to value soft aspects of a job (e.g. good relationships, work-life balance) more and hard aspects (e.g. wage, job security) (Sousa-Poza & Sousa-Poza, 2000, p. 529) less than their male counterparts. However, some studies do not find a significant association between higher income and job satisfaction in Austria, Belgium, Finland, France and Great Britain (Clark et al., 1996, p. 66; Pouliakas & Theodossiou, 2010).

Work time: In economic literature it is assumed that working hours are negatively related to job satisfaction (Bauer, 2004, p. 3). Evidence for this assumption was provided using survey data from the EU (Bauer, 2004), Great Britain (Clark et al., 1996; Asadullah & Fernández, 2008), the USA (Bender, 2005) and the OECD (Clark, 1998). Yet, Clark (1998, p. 12) notes a higher significance of the effect

for women, while Bender finds just the opposite (2005, p. 486) and Asadullah & Fernández (2008) do not do not find any difference between genders.
There are also some contradicting results: For instance, in Denmark and Portugal a positive relationship between working hours and job satisfaction was observed (EFILWC, 2007, p. 18).

Type of work: Several studies find that employees in a management position tend to exhibit higher job satisfaction in Europe (Bauer, 2004, p. 12), Great Britain (Clark et al., 1996, p. 65) and the USA (Near et al., 1978, p. 253).
Furthermore, according to a study by the EFILWC (2007, p. 12) self-employment is positively associated with job satisfaction in Bulgaria, the Czech Republic and the Netherlands. Moreover, Benz & Frey (2008) show that the self-employed receive higher happiness from work than regular employees in Germany, Great Britain and Switzerland.
The reason for these findings might lie in higher task variety, task complexity, autonomy, flexibility and control - factors which all relate positively to job satisfaction (EFILWC, 2007, p. 14, p. 19; Fasang et al., 2007, p. 10 Hooff, Geurts, Beckers, & Kompier, 2011, p. 72).
The role of control and autonomy is among the few topics related to job satisfaction which have been intensively explored for the Japanese context. For instance, Kawada & Otsuka (2011) provide evidence that low job satisfaction of Japanese employees is - among other reasons - related to limited job control. Tokuda et al (2009) confirm that low work control is directly associated with job dissatisfaction among Japanese physicians. Finally, Yamaguchi (2013) finds a positive effect of team autonomy on job satisfaction.
According to results of surveys from Austria, the Czech Republic, Denmark, the Netherlands and Germany, regular workers with permanent employment contracts are more satisfied than those with fixed-term contracts or temporary workers (EFILWC, 2007, p. 11; Graaf-Zijl, 2011). It is argued that the lack of job secu-

rity, but also other unobservable factors are responsible for this gap (Graaf-Zijl, 2011).

Commuting time: For most working people commuting is an important and time-consuming component of their daily routine (Roberts, Hodgson, & Dolan, 2011). Besides, it generates additional out-of-pocket costs (gas, train tickets), causes stress and intervenes in the relationship between work and family (Stutzer & Frey, 2008, p. 341). Yet, there are not much studies which explicitly look at the relation between commuting time and job satisfaction. One recent study by Pfaff (2014) does not find any significant relation, while Steinmetz (2014) suggests that long commuting time negatively affects the intention to stay with the same employer. However, there is a well established link between long commuting times and different areas of life satisfaction. For instance, Stutzer & Frey (2008) and Pfaff (2014) find a statistically significant negative effect of commuting distance on reported life satisfaction and Delmelle et al. (2013) show that one-way commutes of 30 minutes or longer are connected to lower levels of social satisfaction. Since this study controls for income, which is one of the potential variables through which people compensate for long distances between their home and work place (jobs that require longer commutes usually offer higher income), the researcher expects a negative relationship between long commuting time and job satisfaction (Stutzer & Frey, 2008; Pfaff, 2014).

2.2.2 Measures of Subjective Well-being

Measurements of subjective well-being capture satisfaction levels over a wide range of domains, such as satisfaction with one's financial situation or health condition (OECD, 2013, p. 10). Their specific characteristic is *"that only the person under investigation can provide information on their evaluations, emotions and psychological functioning (OECD, 2013, p. 10)."* In respect to job satisfaction - which itself is a measure of subjective well-being - the following links have been observed:

Self-reported health: Earlier studies have shown a positive relationship between self reported health and job satisfaction (Asadullah & Fernández, 2008, p. 25; Clark et al., 1996, p. 65).

Satisfaction with housing: Near et al. (1978, p. 259) find a significant positive relationship between the condition of the respondent's house or neighborhood and job satisfaction. Further, Weaver (Weaver, 1974, p. 374) notes that satisfaction with one's housing situation is positively correlated with job satisfaction.

Self-reported happiness: There is evidence for a reciprocal relationship between job satisfaction and self-reported happiness (Bowling et al., 2010, p. 923). This means that job satisfaction and life satisfaction causally influence each other (Judge & Watanabe, 1993, p. 945). This means that better jobs make people happier, but happy people are also more likely to find their jobs satisfying. However, a number of longitudinal studies have proved that the causal effect is significantly stronger from life satisfaction to job satisfaction (e.g. Headey & Wearing, n.d., p. 114; Judge & Watanabe, 1993, p. 945).

2.3 Effect of Leisure Activities on Job Satisfaction

Leisure is defined as *"those activities which people do in their free time, because they want to, for their own sake, for fun, entertainment, self-improvement, or for goals of their own choosing, but not for material gain (Argyle, 2013, p. 111)."* Though it is frequently assumed that different types of recreation (e.g. meditation, sports, etc.) can improve work life balance and job satisfaction, only few studies have explicitly looked at the link between leisure and job satisfaction. One of them was conducted by Martinez (1994) and examines the relationship between work and leisure within the context of an employee leisure program[6]. The study observes that both job and leisure satisfaction was higher among employees participating in the program. In addition, the frequency of participation was positively

[6] The program included personalized fitness testing, computerized cardiovascular and strength training, aerobic/exercise classes, racquetball, squash, wellness programs, educational programs (single-parent series, self-esteem series) and various recreation programs such as golf, softball, basketball and karate.

related to both leisure and job satisfaction.

A bigger body of research provides several explanations of how leisure satisfaction or specific leisure activities can improve psychological health (Iwasaki, 2001) making it easier to cope with unpleasant sides of working life and thus reduce dissatisfaction with one's job (Pearson, 1998; Pearson, 2008). For instance, certain non-work activities and leisure satisfaction can facilitate:

1. Recovery from work: Sonnentag et al. (2014) indicate that engaging in pleasurable leisure activities can increase psychological detachment from work during off-job time and thus reduce exhaustion. VanHoof et al. (2011) support that pleasurable activities during leisure time have beneficial effects on recovery. Furthermore, Korpela & Kinnunen (2010) identify exercise and being outdoors during free-time as the most effective activities for recovery from work. Oerlemans et al. (2014) show that social and physical activities are positively associated with recovery from work when happiness during such activities was high, but negatively when happiness was low.

2. Coping with stress: Leisure satisfaction and certain leisure activities can reduce stress levels (e.g. Iwasaki, 2001; Yarnal, Almeida, & Qian, 2014). Stanton-Rich & Iso-Ahola (1998) indicate that high leisure satisfaction and specifically leisure activities which promote a sense of self-determination (e.g. playing games, sports and talking to friends)[7] can weaken the negative effects of stress on health, and thus prevent burnout. Furthermore, Mojza & Sonnentag (Mojza & Sonnentag, 2010) find that volunteer work engagement can buffer the effects of job stressors on the subsequent working day.

3. Personal growth: A growing body of research provides evidence that under certain conditions stressful life events (e.g., disrupted relationships, health problems, death of a loved one) can induce positive changes in one's life (e.g. psychological, interpersonal and spiritual growth like an improved self-concept,

[7]Watching TV - offering virtually no challenge, requiring a very low level of skill, and providing the least amount of personal control - was identified as the activity with the least positive effect.

strengthened relationships and deepened spiritual beliefs) (Chun, Lee, Kim, & Heo, 2012). Such changes are subsumed under the term stress-related growth (SRG). Several studies have suggested that certain leisure activities can facilitate SRG. For example, Chun et al. identify frequent participation in civic activities (e.g., volunteering at community organizations or community service) as well as leisure satisfaction in general as statistically significant predictors of SRG (Chun et al., 2012).

4. Higher life satisfaction: As mentioned above job satisfaction is causally affected by life satisfaction. Therefore leisure satisfaction or certain types of recreation could be indirectly linked to job satisfaction through their effect on happiness in general. There is evidence that certain types of leisure activities can bring people into a good mood (e.g. listening to cheerful music or taking a brisk walk) (Argyle, 2013, p. 111). Furthermore Balatsky & Diener (1993) have identified leisure satisfaction as the best predictor for life satisfaction.

3 Data and Econometric Approach

3.1 Data and Sample

The empirical analysis in this thesis is based on micro-data from the large-scale Japanese General Social Survey (JGSS) of the year 2010. The Japanese General Social Surveys are designed and carried out by the JGSS Research Center at Osaka University of Commerce (Joint Usage/Research Center for Japanese General Social Surveys accredited by Minister of Education, Culture, Sports, Science and Technology), in collaboration with the Institute of Social Science at the University of Tokyo. It is designed as a nationally representative household survey covering males and females between the age of 20-89. The total sample size is 9000 and the applied sampling procedure is two-stage stratified random sampling.[8] The survey was conducted from the beginning of February through April 2010 through a combination of interviews and self-administered questionnaires. There are two different self-administered questionnaires: form A and form B. Form A was randomly distributed to half of the participants, and form B was given to the rest of the sample. The respondence rate for form A is 62.18% (n = 2,507) and for form B is 62.14% (n = 2,496). While some survey questions are covered by both forms (e.g. questions on the respondents' level of job satisfaction, occupation type and other basic attributes), other questions (e.g. questions about the respondents' level of happiness in general, the level of solidarity at their workplace and the frequency of some specific leisure activities) only appear on one of the forms.

For the empirical analysis, the following groups were excluded: respondents without a job, employees older than 69 years[9] (due to the small number of participants falling into this category) and those for whom the selected model variables were missing. As a result, our main sample used for the analysis consists of 2209 indi-

[8] More information on the data collection process is available from http://jgss.daishodai.ac.jp/english/surveys/sur_top.html

[9] According to offical OECD statistics the effective age of retirement is usually below the official age for receiving a full old-age pension in most countries. However, Japan and Korea constitute notable exceptions where the effective age of retirement is close to 70 for men in spite an official retirement age of 60 (OECD, 2012).

viduals.

Due to the wide array of information provided in the JGSS,[10] a big range of control variables could be used for the econometric analysis. In the following, a brief description of the dependent and independent variables used in the different models of the empirical analysis is provided.

3.2 Dependent Variable

This thesis examines the relationship between the dependent variable job satisfaction and different social variables, measures of subjective wellbeing and types of leisure activities. Job satisfaction describes the degree of positive feelings towards one's job (Srivastava, 2013, p. 159). In the survey the variable is assessed via question Q13 of the interview ("on the whole, how satisfied are you with the (main) job you have? / genzai no shigoto ni dono kurai manzoku shite imasu ka"), which requests the participants to give an integer rating from 1 (satified / manzoku shite iru) to 5 (dissatisfied / fuman de aru). Figure 1 (p. 19) shows a bar chart for the distribution of the ratings from the main sample. It is apparent at first glance, that the values in this scale are skewed towards the more satisfied end. To make the results of the models more reliable the five categories were changed into 4 by aggregating the respondents who choose 4 or 5 into one category. This is a common practice in academic literature dealing with similarly skewed subjective well-being data (e.g. Oshio, Nozaki, & Kobayashi, 2010, p. 359; Oshio & Kobayashi, 2010, p. 637). Furthermore, the order was reversed - so that higher numerical values represent higher levels of job satisfaction - to simplify the discussion of the estimation results. Figure 2 (p. 20) displays the bar plot for the new distribution:

[10]In total, the survey contains over 560 variables.

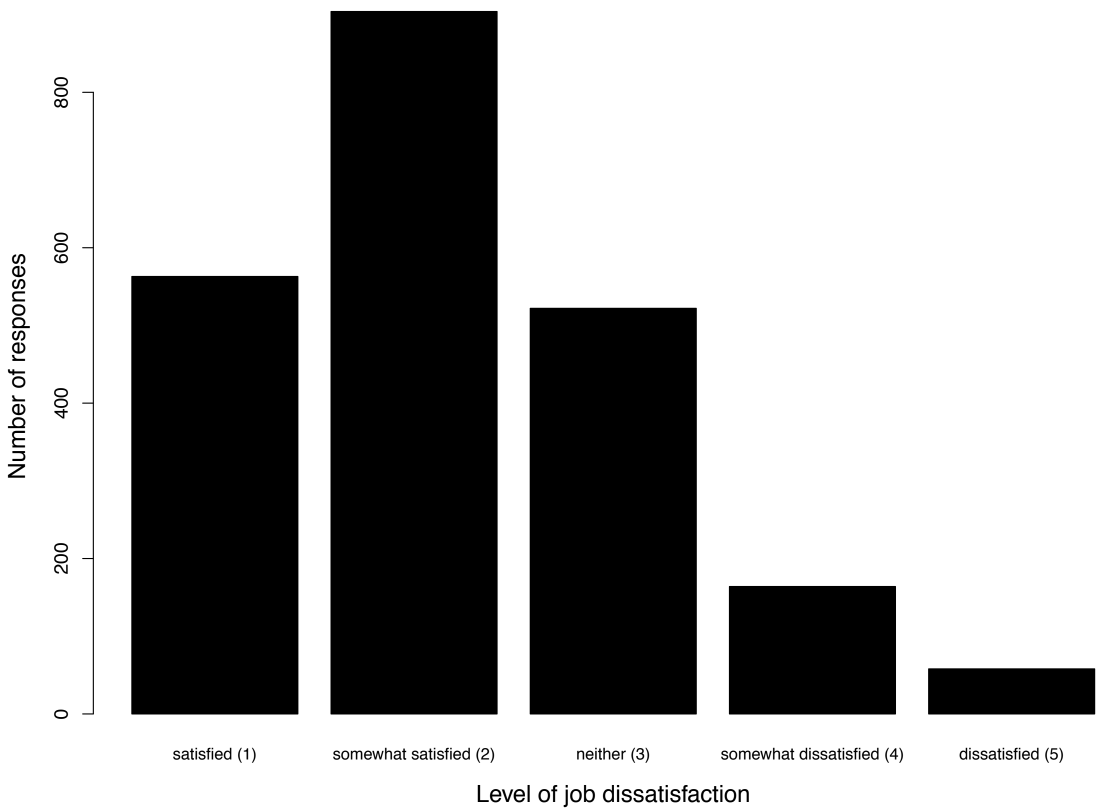


Figure 1: Distribution of original job satisfaction ratings

3.3 Independent Variables

3.3.1 Main Sample

The choice of the independent variables was guided by the findings described in the literature review and covers social and well-being variables commonly appearing in economic studies on job satisfaction. In the following, the variables used in the empirical analysis are briefly described. A more detailed explanation of all variables can be found in Section A ("Definition of all Variables") of the appendix.

1. Age is measured in years. In order to have a sufficient number of respondents in each category[11] employees older than age 69 years were excluded from the sample. This resulted in a age range from 20 to 69 years. *Age*^{2} represents the squared

[11] At least 15 answers for each category.

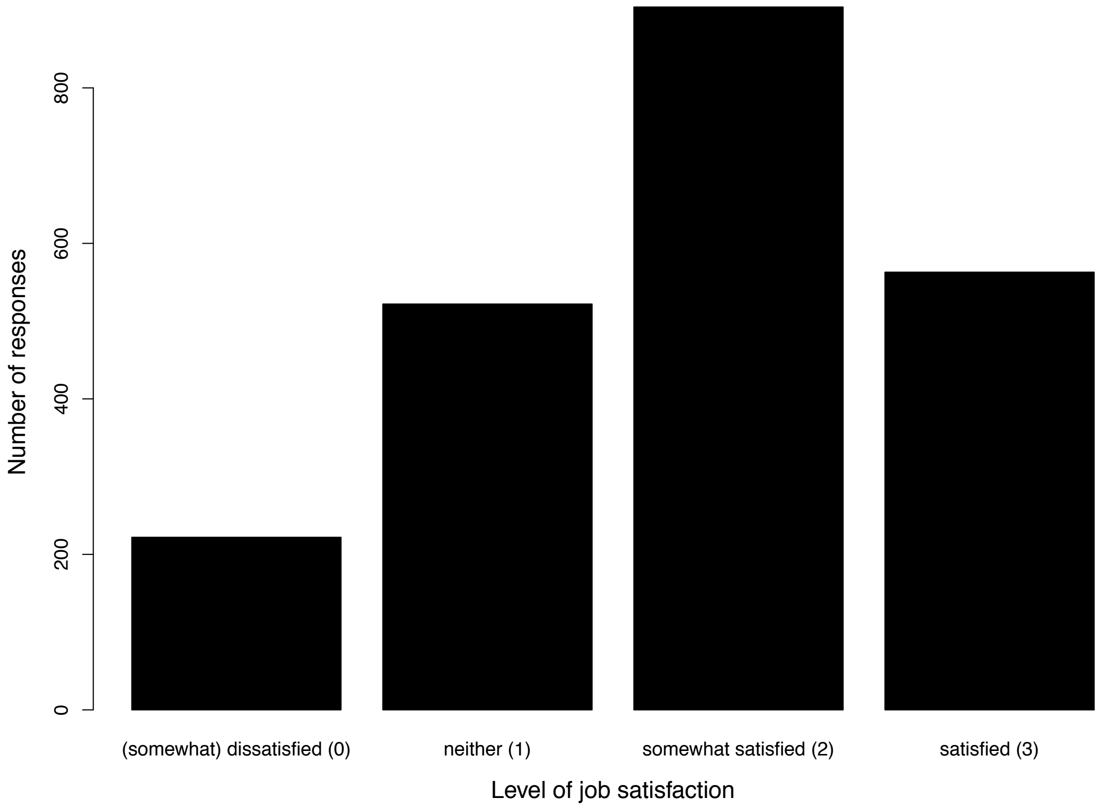


Figure 2: Distribution of final job satisfaction ratings

values of the age variable. Both variables are needed to depict a U-shaped relationship between job satisfaction and age.[12]

2. Gender is coded as a dichotomous dummy variable, with "female" as the reference category.

3. Marital status was coded as a set of dummy variables with "married" as the reference value. Missing values and respondents who answered with "separated" were excluded from the sample due to the lack of observations. The selected categories are "never married," "divorced" and "widowed."

4. Number of children is coded as an ordinal variable with five categories (no, one, two, three, and four and more children).

5. Graduated from college is a dichotomous dummy variable which measures

[12]More information quadratic equations can be found in Studenmund (2005, p. 216).

whether the respondent graduated from college. It is used as a proxy measure for education.

6. Respondent's yearly income is coded as an ordinal variable with thirteen categories. The lowest category represents a yearly income of less than 700,000 ¥,[13] while the highest category represents earnings above 10 million ¥.[14]

7. Work hours is a continuous variable and ranges between 1-112 work hours per week for our sample. Since a negative relationship at a decreasing rate is assumed based on the existing literature - meaning that every additional hour contributes less to job dissatisfaction - the variable is included in terms of its natural log in the empirical analysis.[15] Indeed, for most studies on job satisfaction were work hours are included as a control variable in the regression model the log values are used (e.g. Clark et al., 1996, p. 65; Asadullah & Fernández, 2008).

8. Type of work is a categorical variable which was divided into six categories consisting of "regular employees," "executives of a company or a corporation," "temporary workers" (including daily workers and part-time temporary workers, doing work at home), "dispatched workers from temporary personnel agencies," "self-employed" and "family workers."

9. Employed at government agency can be seen as the seventh category of the *type of work variable*, but is stated here separately, because it is derived from another survey question.

10. Commuting time is coded as an ordinal variable with ten categories. The lowest category covers employees who need less than nine minutes to commute to their work place, while the highest category includes those who need at least 90 minutes.

11. Satisfaction with health condition is measured via question Q12 which appears on both self-administered forms and asks the participants to give a 5-point integer rating going from "dissatisfied" (Jap: "fuman") to "satisfied" (Jap: "manzoku") on

[13]This equals 4.770 € based on the exchange rate from 11th December, 2014.
[14]This equals 68.198 € based on the exchange rate from 11th December, 2014.
[15]More information on the semilog functional form can be found in Studenmund (2005, p. 213).

their health and physical condition. The weakness of this variable is that the question on which it is based does not explicitly request the respondents to give a rating on their health state, but on their level of satisfaction with it. However, it is the only measure of the participants' health condition which applies to the whole sample and is therefore used as a proxy measure for self-reported health.
12. Satisfaction with area of living measures the satisfaction level on the same integer scale as described above.

In addition to the social and well-being determinants above, this thesis aims to also take a closer look at the relationship between job satisfaction and leisure satisfaction as well as different types of non-work activities. Therefore the following variables were also included in the empirical analysis:
13. Satisfaction with leisure provides the respondents' satisfaction levels regarding their non-work activities.
14. Membership in civil or social group is a dichotomous dummy variable indicating whether the respondents are members of either a citizens' movement, a consumers' cooperative or a social service group
15. Membership in hobby organization is a dichotomous dummy variable specifying, if the respondent is part of a hobby group (e.g. photography, mountain hiking, etc.).
16. Frequency of meeting friends captures how often the respondents meet their friends. The lowest value for this ordinal variable is "never" and the highest "several times the a week."
17. Frequency of doing sports measures the frequency in which the respondents engage in exercise or playing sports. The ordinal variable has five categories, of which "scarcely any exercise" constitutes the lowest and "more than several times a week" the highest.
18. Number of books, magazines and comics per months is measured as an interval variable with "zero books" as the lowest category and "four books" as the highest.

19. Hours of television per day is measured by an ordinal variable with eight categories - "zero hours" is the lowest and "seven hours and more" is the highest.

3.3.2 Subsamples based on Questionnaire Responses

As mentioned in the literature review, one possible explanation how leisure satisfaction and the type of recreation affects job satisfaction might be indirectly through its positive influence on overall happiness. Although form A as well as form B include a question on overall happiness, they cannot be combined into one variable for the main regression analysis, since the questions are stated differently leading to clearly distinct distributions of the answers. Therefore and to include other meaningful variables, such as *self-reported health condition* and *solidarity at the participant's work place*, additional regression analyses are conducted on specific variables only accessed in form A and B. Furthermore, some variables, which have not been mentioned in previous studies, but appeared meaningful to the researcher - such as *self-reported social status* and *anxiety regarding economic future* were included in these analyses, although they appeared on both questionnaire forms, to keep the regression models based on the main samples as comparable as possible to previous research studies.

The regression models based on the responses from form A include the following additional variables:
20. Frequency of private trips states how often a respondent goes on a trip which takes more than two days (business trips are excluded). It was included, to further investigate how different forms of leisure use are related to job satisfaction.
21. Self-reported health corresponds better to health variables in previous studies on job satisfaction, since the corresponding question directly asks the participants to rate their health condition on a five point scale raging from "poor" health condition to "good" condition. To avoid multicollinearity *satisfaction with health condition* variable is not used in the models based on the form A responses.
22. Self reported happiness (A) bases on Q13 of form A. The exact wording of

the question is: "Are you happy?" and the respondents could choose their answer on a five point scale going from "unhappy" to "happy".

23. Progression of economic situation indicates whether the respondents financial situation has improved during the last few years. The variable did not appear in previous research, but can be believed to have an effect on job satisfaction for two reasons. First, the variable is likely to be correlated with career opportunities at one's workplace, since the improvement in the economic situation might result from a promotion. Second, even if the change in the financial situation comes from another source (e.g. promotion of spouse, inheritance etc.), it still could have increased the participant's freedom to choose a more interesting (but perhaps less well paid) job. Interestingly, the variable was also measured via form B of the questionnaire, but was not added to the analysis of the main sample, since it's relation to job satisfaction has not been discussed in previous research papers (at least considering the papers the researcher has seen).

24. Believe in opportunities to improve one's life is a variable, which may shed light on the respondent's personality. For example more optimistic individuals are more likely to believe in their ability to improve their life. The variable is based on a question appearing on both questionnaire forms, which asks the participants, if there are enough opportunities in the Japanese society for them to increase the standard of living for their family or for themselves? There are five possible answers ranging from "sufficient" to "not sufficient at all."

25. Self-reported social status: There is no doubt that people prefer to pick jobs, which are well respected in the society (e.g. medical doctors) and in this way increase their social standing. There is also evidence that the perception of a high social standing leads to higher happiness levels (Posel & Casale, 2011). However, the question whether the perception of one's social standing is also related to job satisfaction has not been investigated so far. Again, the question was included on both questionnaire forms.

26. Anxiety regarding economic future is again a variable which can serve as proxy for the respondent's personality, since it can be assumed that people who feel anxious regarding their future financial situation tend to be less optimistic in general too. However, this type of anxiety could also result from low job security or the employing company struggling economically, which might cause workers to worry about future pay cuts or layoffs. Like in the previous cases the corresponding question also appeared on form B.

27. Trust in people and organizations is the last variable which is included in the analysis of the subsample derived from the answers on the questions on form A. The responding question asks the participants whether most people can be trusted giving three possible answers; No, depends and yes.

There are also three regression models (general, male and female sample) estimated, which are based on the responses from form B. They include partly the same additional variables mentioned in the section above - namely progression of economic situation, believe in opportunities to improve one's life, self-reported social status and anxiety regarding economic future. However, a number of unique variables based on the questions from form B were added:

28. Loneliness: Loneliness is a strong predictor for self-reported happiness (Tiefenbach & Kohlbacher, 2012, p. 27). However, there have not been much, if any research whether it is also connected with job satisfaction.

29. Pain interfering with daily activities is an alternative health measures which focuses on a specific side effect of a poor health condition.

30. Self reported happiness (B) is stated separately because the exact wording of the question differs from the happiness question on form A, which leads to a very different frequency of the different ratings. More specifically, the question on the B form reads as "in general, are you happy?"/"zentai toshite, anata ha, genzai shiawase desu ka" - in contrast to just "are you happy"/"anata ha, genzai shiawase desu ka." Also, the wording for the possible answers on the five point scale are stated differently. While on form A the possible answers run from "happy/shiawase" to "unhappy/fushiawase", it is "very happy/hijou ni shiawase" and "very unhappy/hijou ni, fushiawase" on form B.

31. Solidarity at work place is the last variable which is included and captures the quality of relationships between the respondent and his coworkers. It is assumed that people under the same work conditions, will feel more content with their work place, if they enjoy the company of their coworkers and/or feel that they can rely on them.

3.4 Choice of Method

The dependent variable *job satisfaction* is a latent continuous variable. This means that we do not know its true value, but only observe when it crosses certain thresholds - e.g. from "somewhat satisfied" to "satisfied." Therefore the variable needs to be treated as an ordinal variable.[16] In this case, ordered logistic (or ordered logit) regression[17] is preferable to ordinary least square (OLS) regression for estimating models.[18] The researcher chose the article *"A Tutorial in Logistic Regression"* by DeMaris (2012) as the theoretical foundation for the analysis, because the article provides a comprehensive introduction into major uses of ordered logistic regression in social data analysis. Furthermore, DeMaris uses *happiness*

[16] *Continuous variables* can take any value between their minimum value and their maximum value (e.g. age, height, etc.). In contrast, *ordinal variables* can move only between thresholds, meaning that the number of values it can take is limited (e.g. level of satisfaction with one's life, health or leisure rated on a 5-point scale). This is also true for *interval variables*, which only differ from ordinal variables in the point that the intervals between the values are equally spaced (e.g. number of books per month, number of children, etc.,). *Categorical variables* are similar to ordinal variables, except for the limitation that they do not have a clear ordering between the categories (e.g. occupational status, gender, etc.,).

[17] *"Together with the probit model, the logit model belongs to the class of probability models that determine discrete probabilities over a limited number of possible outcomes"* (Cramer, 2003, p. 2). Since, probit models do not respect the limited range from 0 to 1 which is imposed on probabilities, logit models are preferably used by many researchers (Studenmund, 2005, p. 449*ff*, DeMaris, 2012, p. 957). Another advantage is that the behavior of a logistic function follows a sigmoid curve, which serves as a good approximation for binary real-world data (Cramer, 2003, p. 12; Studenmund, 2005, p. 455).

[18] OLS is the most widely used regression method in social sciences and derives estimates by minimizing the sum of squared residuals between observed and predicted values. However, OLS requires the dependent variable to be cardinal and does not account for the variable to be bounded (D'Addio et al., 2007, p. 2415), like in the case of the dependent *job satisfaction* variable used in this thesis.

as the dependent variable in his equations, which has similar properties as the dependent variable *job satisfaction* used in this thesis.

Although, OLS regression is in theory a less appropriate method of estimation in the present case, the researcher decided to provide the OLS estimates as well, for several reasons:

1. Interpretation: OLS coefficients are more readily interpretable than logit coefficients, which first would need to be translated into "log-odds ratios," a relatively unknown concept which would need to be introduced first (e.g. Kalleberg & Vaisey, 2005, p. 441; Eckey, Kosfeld, & Dreger, 2011, p. 154).

2. Comparability: Many psychological studies on job satisfaction use OLS regression instead of more appropriate alternatives, such as ordered logistic or probit regression (e. g. Kalleberg & Vaisey, 2005; Bianchi, 2012).[19] Providing additional OLS estimates helps to put our results into context, since it makes findings more comparable to previous and future research and shows whether the identified effects are persistent no matter which estimation method is used.

3. Cardinality: One reason economists avoid the use of OLS on well-being data, is that it is based on the implicit assumption that the dependent variable is cardinal, meaning that the self-reported satisfaction ratings are equidistant (Luttner, 2005, p. 981; D'Addio et al., 2007, p. 2415; Gandelman & Porzecanski, 2011, p. 259). However, Ng (1997) gives several arguments why measures of subjective well-being can be treated as cardinal. Indeed, a growing body of research provides evidence, that OLS regression used on subjective well-being data often leads to similar results considering significance and direction of the coefficients than ordered logit as ordered probit models (e.g. Ferrer-Carbonell & Frijters, 2004). By providing both types of coefficients - those estimated through OLS regression and ordered

[19] In many of these studies the robustness of the results were verified by estimating an ordered logistic and probit regression and comparing whether these models identified the same significance and direction of the estimates as the OLS models (f. e. Kalleberg & Vaisey, 2005, p. 441; (Bianchi, 2012, p. 279)). However, these studies usually did not provide regression tables with the original results from the logit or probit models.

logit regression - this thesis adds to the methodological literature dealing with the comparability of these two approaches in regard to data on job satisfaction.

4. Enablement of advanced statistical techniques: In order to take into account different factors which might bias the results (e.g. endogeneity, heteroskedasticity, multicollinearity or irrelevant variables) some advanced statistical tools need to be applied. However, some of them are only available for OLS models, since it is the most common regression technique. In the case of this thesis methods to account for endogeneity and irrelevant variables were only available for OLS models, while other statistical tools - e.g. robust standard errors (to deal with heteroskedasticity) and variance inflation factors (to test for multicollinearity) - could fortunately be obtained for both types of models.

3.5 Model Estimation

The index model for the latent continuous variable *job satisfaction* (js^*) for individual i is given by:

$$js_i^* = X_i'\beta + \varepsilon_i \tag{1}$$

where X_i represents a vector of the independent explanatory variables for individual i, β the parameter vector to be estimated and, ε_i the error term. Since the continuous latent variable js^* cannot be observed, it is replaced by the ordered categorical response variable *self-reported job satisfaction* js with k = 0, ..., 3 categories[20] and individual-specific thresholds λ_k^i, where $\lambda_k^i < \lambda_{k+1}^i$:

$$js_i = k \iff \lambda_k^i \leq js_i^* < \lambda_{k+1}^i \tag{2}$$

The estimates of the parameter vectors β are obtained via ordered logistic regression. In addition, the parameter vector is also estimated using OLS regression treating job satisfaction as a continuous variable.

All parts of the data analysis were conducted using the *"R Project for Statistical Computing,"* a free but powerful software environment for statistical computing

[20] The categories are "dissatisfied or somewhat dissatisfied," "neither satisfied nor dissatisfied," "somewhat satisfied" and "satisfied."

and graphics available from http://www.r-project.org. The source code, including comments, on the data analysis can be found in Section F of the appendix.

3.6 Additional Considerations

Prior to the estimation procedure some statistical issues need to be considered:

3.6.1 Causal Direction and Endogeneity

Determining the direction of causality is a crucial step in data analysis, because only if a causal link from the independent variable to the dependent variable can be established, a theoretical foundation for the attempt to increase satisfaction via policy intervention exists (Frey, 2008, p. 11).

Commonly, the following three conditions must be met before a researcher can claim that x causally affects y (e.g. Kenny, 1979, p. 3; Antonakis, Bendahan, Jacquart, & Lalive, 2014, p. 5)

1. Time precedence: x precedes y in time.

2. Relationship: A significant relationship between x and y is present.

3. Nonspuriousness: There is no correlation between x and the error term. This means that there are no unmodeled causes, which would eliminate or invert the statistical relationship between x and y, if they had been included.

The following subsection will specifically deal with the last condition, or more specifically what happens if it is not fulfilled.[21] In such a case x is said to be *endogenous*. This means that in the worst case, the coefficient estimates of the endogenous variable as well as other variables of the model might be biased (Antonakis et al., 2014, p. 6). More specifically, if no measures are taken to account for the endogenous variables, estimated coefficients of all variables in the model are at risk to be inconsistent and thus misrepresent the true relationships between the dependent and independent variables.

[21]In the subsection "Sources of Endogeneity" in Section E (Statistical Tools and Concepts) of the appendix different causes for the violation of this condition are listed. In addition, it is described how they specifically apply to the JGSS data.

Depending on the type of data there are different methods to account for endogeneity:

1. In *experimental studies* endogeneity can be minimized through *experimental design*, by selecting participants with similar characteristics and then randomly allocating them into a treatment or control group (Stock & Watson, 2006, p. 85). Under this condition the causal effect of the experimental treatment can be measured as the difference between the treatment and control group (Antonakis et al., 2014, p. 13).

2. *Panel data* is defined as data containing observations of the same variables over multiple time periods for the same individuals, institutions or countries. Panel data allow the explicit *modeling of omitted fixed effects* (for example personality traits) to mitigate the effect of endogeneity (e.g. Wooldridge, 2002, p. 441*ff*, Antonakis, Bendahan, Jacquart, & Lalive, 2010, p. 1092).

3. However, in the case of *cross-sectional data* the means to account for endogeneity are more limited. Cross-sectional data is for example data that includes observations of many different subjects at the same point of time. Unfortunately, for this study only this type of data is available. Considering the means to account for endogeneity the probably *"most useful and most-used method"* (Antonakis et al., 2010, p. 1100) is two-stage least squares (2SLS) regression (also known as instrumental-variable (IV) regression), which requires at least one instrumental variable for each endogenous variable that satisfies the following properties (Wooldridge, 2002, p. 563).

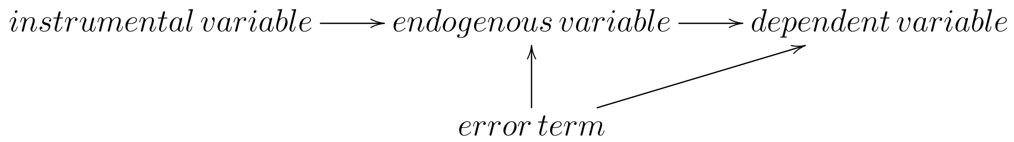
a. It is strongly correlated with the endogenous variables for which it was introduced as an instrumental variable.

b. It is uncorrelated with the error term.

c. It is has no effect on the dependent variable y (except the indirect route via the endogenous variable in question).

This conditions can be visualized by the following diagram (*4.8 Instrumental*

Variables, n.d., p. 37):



The 2SLS procedure requires the endogenous variables to be replaced by estimates, which were generated by regressing the endogenous variable on their instrument variable as well as the exogenous variables (variables uncorrelated with the error term) in the original equation of the model (Katchova, 2013, p. 6). This allows to obtain consistent coefficients and solve the endogeneity problem, since the estimates are derived from purely exogenous variables and thus free from the original (endogenous) variable's components causing endogeneity.

The researcher tried out several approaches to account for endogeneity by using this method, but has come to the conclusion that for the models of this thesis it is probably an impossible task.[22] There are two reason for this evaluation.

First, the *"goodness-of-fit"* indicators R^2 and *MacFadden's pseudo* R^2[23] have very low values,[24] meaning that a substantial degree of variation remains unexplained. This fact makes it very difficult to find variables which satisfy the second condition (no correlation with the error term). For instance, the researcher considered the use of participation frequency in certain types of leisure activities - namely doing sports and seeing friends - as instrumental variables for leisure satisfaction,

[22]Indeed, dealing with endoneity in studies on job satisfaction is a very challenging task, and thus avoided by most researchers unless they use experimental or panel data for their estimation. No single cross-sectional job satisfaction study, which the researcher encountered, has accounted for endogeneity, for example by applying instrumental-variable regression. Some of the studies explicitly state that the treatment of the endogeneity bias is not possible due to the cross-sectional character of the data and thus refrain from any speculation about causal realationships (e.g. Sousa-Poza & Sousa-Poza, 2000, p. 521), while most cross-sectional studies do not even mention the endogeneity problem (e.g. Clark et al., 1996; Clark, 1997; Nielsen & Smyth, 2008).

[23]A definition of these measures can be found in Section E (Statistical Tools and Concepts) of the appendix.

[24]A low R^2 is not unusual in happiness studies (e.g. Hanson & Sloane, 1992; Loscocco & Bose, 1998, p. 105; Sousa-Poza & Sousa-Poza, 2000, p. 528), because subjective well-being strongly depends on unobservable personality traits (e.g. Veenhoven, 1994, p. 11; Roelen, Koopmans, & Groothoff, 2008, p. 433; Graaf-Zijl, 2011, p. 204).

since they are good predictors of leisure satisfaction, but have insignificant, close to zero coefficients as regressors for job satisfaction. However, additional research later revealed that personality and leisure activity preferences - as well the degree of participation - are significantly related (e.g. Howard, 1976; Lu & Kao, 2009, p. 3). In particular, extraversion - a personality variable which is expected to be associated with job satisfaction (e.g. Lounsbury, Moffitt, Gibson, Drost, & Stevens, 2007; Zhai, Willis, O'Shea, Zhai, & Yang, 2013), leisure satisfaction (e.g. Lu & Hu, 2005; Lu & Kao, 2009) and happiness in general (e.g. Veenhoven, 1997, p. 15; Hills & Argyle, 2001; Frey, 2008, p. 24) - are correlated with the participation in various sports and social activities (Furnham, 1981). This is problematic in respect to this study, because it means that the instrument variables which the researcher had in mind most probably are related to the error term.

In a second attempt to cope with the endogeneity problem, the researcher used a method developed by Lewbel (2012), which uses heteroscedastic covariance restrictions to obtain unbiased estimates. Unfortunately, the instruments generated by this method were too weak and thus could not be used.[25]

The second problem arises from the difficulty to determine which of our model's variables should be treated as endogenous and which as exogenous. The subjective well-being variables are clearly endogenous while age and gender are clearly exogenous, but what is about marital status or employment type? Could they not also be related to omitted variables - such as psychological traits and thus be correlated with the error term?

[25] An instrument variable is considered to be weak, if it's association with the endogenous variable is not strong enough to mitigate the endogeneity bias. This can lead to wrongly identified causal effects (Bound, Jaeger, & Baker, 1995). A widely used method to identify such variables is the first stage F-statistic for joint significance of instruments (Antonakis et al., 2014, p. 32). Based on the findings by Stock, Wright & Yogo (2002, p. 522) many researchers stick to a rule of thumb classifying instruments as weak, if the first stage F-statistic falls below the threshold of 10 (e.g. Stock & Watson, 2006, p. 441; Angrist & Pischke, 2009, p. 157; Baltagi, 2011, p. 267). Unfortunately, in the case for our models the F-statistic never exceeded the threshold of 10.

In conclusion it can be said that the high number of possible sources[26] for endogeneity made the application of the 2SLS procedure extremely difficult, if not impossible. Therefore the researcher needs to refrain from causal claims in this thesis, except in cases where the direction of the effect has been indicated by previous studies based on experimental or panel data. For example, the study by D'Addio, Eriksson & Frijters (2007), which applies a fixed effects ordered logit model, suggests that hourly wage, being a full-time worker and a good health condition are indeed causally related to higher job satisfaction for men, but not necessarily for women.

Even though, we need to be cautious considering causal claims, the findings are still meaningful, because they might shed light on significant correlations, which can later be investigated in detail by other studies using panel and experimental data to account for endogeneity.

3.6.2 Heteroscedasticity

Unlike endogeneity, heteroscedasticity does not affect the consistency of the coefficient estimates, but instead leads to biased standard errors (SEs), which cause over or understated p-values (Antonakis et al., 2010, p. 1098). In the case of our analysis, the Breusch–Pagan test[27] indicates the presence of heteroskedasticity in some of the OLS as well as the ordered logit models. Fortunately, the bias can be easily resolved, since robust standard errors (also known as Huber-White standard errors or sandwiched standard errors) (Antonakis et al., 2010, p. 1098) can be easily obtained for both types of models. Note, that for all models presented in this thesis robust standard errors have been estimated and displayed in the regression tables. However, the robust SEs in all models of this thesis only slightly deviate from the traditional, potentially biased SEs. This also leads to almost the same significance levels, no matter whether robust or traditional SEs

[26] Possible sources for endogeneity in respect to the models of this thesis are listed in Section E (Statistical Tools and Sources) of the appendix.

[27] This test is commonly used for cross-sectional data (Studenmund, 2005, p. 361). For a detailed explanation see Wooldridge (2002, p. 257*ff*) or Eckey et al. (2011, p. 90*ff*)

are used. This means that though heteroscedasticity might be present in some of the initial models, it does not bias the SE estimates or the significance levels to a large degree.

3.6.3 Multicollinearity

Multicollinearity occurs when there are linear dependencies among explanatory variables. It can become a serious problem, if the degree of correlation among explanatory variables is too high (Curto & Pinto, 2011, p. 1499). More specifically, if two or more explanatory variables are highly correlated with one another, it becomes very difficult to separate their distinct impacts on the dependent variable as long as all of them remain in the model (P. D. Allison, 2012, p. 60). Multicollinearity does not directly bias coefficients, but it can cause instability by over-inflating standard errors leading to the underestimation of significance values (P. D. Allison, 2012, p. 60). In the worst case scenario, variables which are in fact significant may appear as insignificant due to the presence of multicollinearity (Martz, 2013, p. 24). Fortunately, these consequences only apply to those variables that are collinear (P. D. Allison, 2012, p. 60).
There are various ways to measure the magnitude of multicollinearity and detect problematic variables. Observing the correlation matrix helps, but is not sufficient, since it only captures the correlation between two variables, but not the interdependence of several variables as a group (P. D. Allison, 2012, p. 60). A better method to quantify the severity of multicollinearity and identify the responsible variables is to estimate a measure called variance inflation factor (VIF) (Martz, 2013, p. 24). The VIF for a specific independent variable x is calculated using the coefficient of determination (R_x^2) resulting from regressing this specific variable on the other explanatory variables of the original model (Curto & Pinto, 2011, p. 1500):

$$VIF_x = \frac{1}{1 - R_x^2} \tag{3}$$

The interpretation of the measure is straightforward; the higher the value, *"the more troublesome the variable"* (Curto & Pinto, 2011, p. 1500). Since VIFs increase dramatically when the belonging R^2 exceeds 0.9, a rule of thumb is to remove variables with VIF values larger than 10 (Curto & Pinto, 2011, p. 1500) from the model. Another rule of thumb sets the critical value lower and considers multicollinearity as problematic, if the VIF exceeds 5 (Martz, 2013, p. 25).

However, there are some exemptions from these rules, such as in the case of intentionally included powers or products of other variables: If a regression model includes both x and x^2 - like in our models *age* and *age*^{2} - these two variables most likely will be highly correlated. Similarly, if a model has x, z, and xz, both x and z will be highly correlated with their product. However, this relationships are not problematic, because the correlation does not affect the significance levels (P. Allison, 2012).[28]

Therefore, multicollinearity is treated in this thesis in the following way: Except in the case of *age* and *age*^{2}, variables with a VIF higher than 10 are step-wise removed from the model, starting with the variable with the highest score. The VIFs are displayed in every regression table in square brackets right below the standard errors of the coefficient values.

3.6.4 Missing Values (Income Variable)

The number of missing values is really small for most variables, therefore *listwise deletion*[29] is not problematic in most cases (Lynch, 2003, p. 9). The only exception is the variable *respondent's income*, for which a big fraction of values is missing. More specifically, there are 494 missing values, which constitute approximately 18% of the total sample. Since we cannot assume that the income data

[28]This is can be easily demonstrated by reducing the correlation by "centering" the variables (e.g. subtracting their means) before creating the powers or the products. The result will be the same p-values for x^2 or for xz as in a model without this modification, proving that multicollinearity has no adverse consequences for these specific cases (P. Allison, 2012).

[29]Listwise deletion is a method for handling missing data in statistics. In this method, an entire record is excluded from analysis, if any single value is missing.

is missing at random (MAR),[30] standard treatments of missing values - such as *listwise deletion* or *mean imputation* are likely to create biased results (Lynch, 2003, p. 3). Even *Heckman Selection Modeling* - which usually would be the method of choice to handle non-MAR missing values (Lynch, 2003, p. 4) - cannot be applied in this case, because income data violates the normality assumption due to a heavier tail (Marchenko & Genton, 2012, p. 304).

Due to its theoretical importance, the variable needs to remain in the main analysis. However, there is the possibility to estimate an alternative model were the problematic variable is replaced with a proxy and compare its results to the original model. In this thesis, three[31] ordered logit models and three OLS models in which the variable *respondent's income* was replaced by the proxy variable *satisfaction with the financial situation of one's household*[32] were estimated.

3.6.5 Irrelevant Variables

Irrelevant variables are variables which have been mistakenly included in the model equation. In some way they are the opposite of omitted variables (Studenmund, 2005, p. 170). Irrelevant variables do not cause bias, but decrease precision by increasing the variances of the estimated coefficients for the explanatory variables (Studenmund, 2005, p. 170). The models in this thesis include a number of variables, which have not been investigated in combination with job satisfaction before (e.g. different types of leisure activities and commuting time) or have often be found to have no relation to job satisfaction in previous studies (e.g. number of children). So it is likely that some of the included variables will be irrelevant.

[30]Data on variable y is missing at random (MAR), if the missingness is related to a particular variable, but not to the value of the variable that has missing data. Phrased in another way, if the missingness on income is also a function of income (e.g. persons with high income do not report their income), then the data is not MAR (Lynch, 2003, p. 1).

[31]One model covers the whole sample while the other two cover either the male or the female sample.

[32]This variable is derived from question Q12 (interview) of the JGSS and asks the respondent to rate "the current financial situation" of his or her household on a five-point-integer scale. The variable contains only 6 missing values which constitute only 0.22% of the total sample. Therefore listwise deletion can be applied (Lynch, 2003, p. 9).

To account for this problem, variables which lead to a lower adjusted R^2[33] when included in the model are stepwise removed from the regression models (starting with the variables with the lowest significance levels) after estimating the first ordered logit, OLS and proxy models for subsequent regression models.

[33]Like the coefficient of determination (R^2), the adjusted R^2 or R-bar-squared ($\bar{R}^2$) measures the goodness of fit for a model. The main difference between both measures is that $\bar{R}^2$ is also designed to adjust for degrees of freedom (Studenmund, 2005, p. 59). In this way it's value only increases, if the improvement in fit resulting from the addition of the new variable offsets the loss of the degree of freedom. In contrast, R^2 always increases, if a new variable is added to the equation - no matter whether or not it is relevant to the model. Therefore, researchers automatically use $\bar{R}^2$ to evaluate relevance of a new variable. The question, which serves as a criterion, is *does the overall fit of the equation (adjusted for degrees of freedom) improve when the variable is added to the equation? (Studenmund, 2005, p. 173).*

4 Results

This chapter presents the results of the data analysis. Since the number of estimated models is relatively high due to the need to account for different statistical problems, it would go beyond the page count to display and discuss all of them in detail in the main body of the thesis. Instead, the chapter concentrates on the models covering the main sample. However, after the discussion of these main models, the results from the additional models are briefly discussed.

The appendix includes detailed definitions of all variables (Section A), the descriptive statistics for all samples (means and standard errors of all explanatory variables) (Section B) and all regression tables (except for the first two of the main analysis, which are displayed in this chapter) (Section C). Furthermore, the correlation matrix for the models displayed in this chapter can be found in Section D of the appendix. Section E contains more detailed information about statistical tools and concepts used in the data analysis and Section F displays the source code for the estimation of all models. For the statistical analysis the free software environment *"R Project for Statistical Computing"* was used.

4.1 Regression Results (Main Models)

Table 1 (OLS) and table 2 (ordered logit) show the regression results based on the main sample. Note that the regression was not only ran for the whole sample, but also separately for women and men. The values without brackets represent the estimated coefficients, values in ordinary brackets are the SEs, and values in squared brackets are the VIFs. For all models robust SEs were estimated.

The regression tables for the OLS models - including those displayed in the appendix - contain the following additional information: Number of observations, the coefficient of determination (R^2), the coefficient of determination adjusted for degrees of freedom ($\bar{R}^2$), the residual standard error, and the F statistic.[34] Furthermore, the regression tables for the ordered logit models contain the following

[34]The F statistic indicates whether a group of variables are jointly significant.

relevant additional measures: Number of observations, the McFadden R^2,[35] the χ^2 value[36] and the log-likelihood.
The significance levels (two-tailed) of the estimated coefficients, the F statistic and χ^2 are indicated in the following way: P-values lower than 0.01 (1% significance level) are marked with three stars (***); p-values higher than 0.01, but lower than 0.05 (5% significance level) are marked with two stars (**); finally p-values in between 0.05 and 0.1 (10% significance level) are market with one star (*).

4.1.1 Results from Main OLS Model

The regression results for the general sample (displayed in table 1) identify the variables *income, work hours (log), satisfaction with health, satisfaction with the area of living, satisfaction with leisure*, the both dummy variables *male* and *never married* and surprisingly the *number of books, comics or magazines* read per month as the determinants with the highest significance (highly significant at a 1% level). As significant at a 5% level the table states *age, age*^{2} - indicating a U-shaped relationship - and *membership in a civil or social organization.*
The variables *never married* and *male* have a quantitatively similar effect on job satisfaction: Being male or not have been married are both associated with an on average 0.25 points lower job satisfaction score (on a scale from 0-3) than their reference values. In contrast, being divorced or widowed seems not to be significantly related to job satisfaction for the general sample. Similarly, the number of children and the commuting time seems not to have a major effect on job satisfaction for the general as well as for the female and male sample. Moreover, the type of employment seems not to affect job satisfaction to a high degree: Among the six considered variables (executive position, temporary employed, dispatched worker, self-employed, family worker and employed at a government agency) only being a dispatched worker or working at a government agency are significant at 10% level.

[35] A detailed explanation of the McFadden R^2 can be found in Section E of the appendix.
[36] Like the F test in OLS-regression, the χ^2 statistic indicates, if a predictor set for a logistic regression model is jointly significant. For the corresponding likelihood ratio test see Engle, 1984, p. 780.

Considering the magnitude, being a dispatched worker is associated with about 0.26 points lower job satisfaction, while government employes exhibit on average about 0.11 points higher job satisfaction than otherwise comparable survey participants. Having graduated from college is mildly significant at a 10% level and associated with an approximately 0.08 increase in job satisfaction. However, the variable is not significant, if only the male or the female sample are considered. The same is true for the variable membership in civil or cocial organization, which is also only significant in the regression model run on the general sample, but not in those run on the male or female sample.

There are also major differences between the gender groups. For instance, *work hours (log)* and *satisfaction with area of living* are not significant for the male sample, but highly significant[37] for the female sample. In contrast, the variables *widowed, family worker* and *number of books/comic/magazines* have a positive highly significant relationship with job satisfaction in the model based on the male sample, but not regarding the female sample. Even in cases where the significance levels are similar, the magnitude of the effect can differ tremendously based on the gender of the participant. For example the variables *never married, income* and *satisfaction with leisure* are highly significant in all models displayed in table 1, yet their magnitude is quite different, if we compare the female with the male survey participants: On average woman, who have never been married are about 0.35 points less satisfied with their job than other comparable female participants, while never married men are only 0.18 points less satisfied than otherwise comparable men; Moving up one income category (of overall thirteen categories) leads to an approximately 0.07 points increase in job satisfaction for men, while women move up only by 0.04 points; Finally, if men experience a one point increase in their leisure satisfaction (on a scale from 0-4), the data suggest an 0.13 increase in their job satisfaction, while job satisfaction for women only rises by 0.07 points under the same conditions.

[37]Highly signifcant means signficant at a 5% level.

Table 1: Regression Analysis Main Sample (OLS)

	general	women	men
age	-0.030^{**}	-0.033^{*}	-0.032^{*}
	(0.012)	(0.017)	(0.019)
	[67.120]	[65.829]	[82.973]
age^2	0.0003^{**}	0.0004^{**}	0.0004^{*}
	(0.0001)	(0.0002)	(0.0002)
	[64.225]	[62.283]	[81.817]
male	-0.251^{***}		
	(0.048)		
	[1.639]		
never married	-0.252^{***}	-0.348^{***}	-0.183^{**}
	(0.068)	(0.102)	(0.093)
	[2.169]	[2.427]	[2.102]
divorced	−0.031	−0.140	0.135
	(0.096)	(0.118)	(0.165)
	[1.065]	[1.099]	[1.045]
widowed	0.083	−0.107	0.549^{***}
	(0.124)	(0.148)	(0.187)
	[1.069]	[1.122]	[1.032]
number of children	0.006	0.003	0.012
	(0.024)	(0.036)	(0.033)
	[2.120]	[2.341]	[2.024]
graduated from college	0.075^{*}	0.073	0.064
	(0.040)	(0.056)	(0.055)
	[1.174]	[1.191]	[1.197]
income	0.058^{***}	0.041^{**}	0.070^{***}
	(0.010)	(0.016)	(0.013)
	[2.629]	[2.078]	[1.951]
work hours (log)	-0.120^{***}	-0.136^{**}	−0.084
	(0.045)	(0.059)	(0.073)
	[1.666]	[1.617]	[1.405]

executive position	0.042	−0.157	0.067
	(0.097)	(0.221)	(0.108)
	[1.167]	[1.113]	[1.214]
temporary employed	0.006	−0.090	0.059
	(0.063)	(0.082)	(0.113)
	[2.130]	[2.261]	[1.561]
dispatched worker	−0.259*	−0.224	−0.436*
	(0.139)	(0.161)	(0.256)
	[1.076]	[1.124]	[1.051]
self-employed	0.003	0.018	−0.005
	(0.075)	(0.129)	(0.091)
	[1.258]	[1.330]	[1.240]
family worker	0.044	−0.145	0.414**
	(0.117)	(0.146)	(0.175)
	[1.163]	[1.300]	[1.075]
employed at government agency	0.114*	0.071	0.144
	(0.064)	(0.090)	(0.092)
	[1.096]	[1.108]	[1.105]
commuting time	−0.002	0.013	−0.008
	(0.009)	(0.015)	(0.011)
	[1.290]	[1.416]	[1.203]
satisfaction with health	0.069***	0.074**	0.058*
	(0.022)	(0.031)	(0.030)
	[1.254]	[1.243]	[1.280]
satisfaction with area of living	0.059***	0.0934***	0.032
	(0.022)	(0.030)	(0.031)
	[1.251]	[1.278]	[1.254]
satisfaction with leisure	0.099***	0.069**	0.132***
	(0.023)	(0.034)	(0.031)
	[1.387]	[1.422]	[1.397]
member in civil/social organization	0.145**	0.159	0.143
	(0.067)	(0.101)	(0.089)
	[1.076]	[1.078]	[1.083]

member in hobby organization	0.080	0.076	0.101
	(0.058)	(0.079)	(0.085)
	[1.082]	[1.132]	[1.067]
frequency of meeting friends	0.010	0.010	0.005
	(0.018)	(0.027)	(0.024)
	[1.206]	[1.196]	[1.250]
frequency of doing sports	0.011	0.008	0.014
	(0.012)	(0.017)	(0.017)
	[1.108]	[1.094]	[1.134]
number of books/comics/magazines	0.055***	0.034	0.072***
	(0.017)	(0.025)	(0.024)
	[1.160]	[1.163]	[1.199]
hours watching TV	−0.007	−0.008	−0.005
	(0.013)	(0.017)	(0.018)
	[1.136]	[1.134]	[1.174]
constant	3.093***	5.108***	−0.071
	(1.095)	(1.541)	(1.578)
observations	2209	1064	1145
R^2	0.112	0.092	0.145
adjusted R^2	0.102	0.070	0.126
residual Std. Error	0.879	0.878	0.880
F Statistic	10.62*** (df = 26; 2182)	4.205*** (df = 25; 1038)	7.567*** (df = 25; 1119)
Note:			*p<0.1; **p<0.05; ***p<0.01

4.1.2 Results from Main Ordered Logit Model

Overall the results from the ordered logit models are consistent and very similar to the findings above. However, there are some differences to the OLS models. In addition to the variables that have been already identified as highly significant at a 1% level in the OLS models[38] age and age^2 also fall into this category in

[38]Namely *male, never married, income, work hours (log), number of books/comics/magazines, satisfaction with health, satisfaction with the area of living* and *satisfaction with leisure.*

the ordered logit models. Furthermore, working in a government agency is not significant (in the OLS model mildly significant at a 10% level) and being member in civil or social organization is only significant at a 10% level (in the OLS model it is 5%) when using ordered logit regression on the general sample.

Considering the female sample there are the following differences: The income variable is highly significant at 1% level (in the OLS model it is 5%) and the leisure satisfaction variable is only mildly significant at a 10% level (in the OLS model it is 5%).

Finally, significance levels for the male sample differ in the subsequent ways for the logit model: Both age and age^2 are highly significant at a 5% level (in the OLS model they are only significant at a 10% level); the dummy variables *never marrried* and *widowed* are only significant at a 10% level (in the OLS model it is between 1% and 5% percent); being a dispatched worker is not significant (in the OLS model the variable is mildly significant at a 10% level) and the satisfaction with one's health is highly significant at a 5% level (in the OLS model it is only 10%).

Table 2: Regression Analysis Main Sample (Ordered Logit)

	general	women	men
age	-0.073^{***} (0.026) [68.095]	-0.082^{**} (0.037) [63.909]	-0.081^{**} (0.041) [93.120]
age^2	0.001^{***} (0.0003) [65.946]	0.001^{**} (0.0004) [60.861]	0.001^{**} (0.0004) [94.390]
male	-0.526^{***} (0.103) [1.743]		
never married	-0.533^{***} (0.141) [2.289]	-0.793^{***} (0.220) [2.734]	-0.352^{*} (0.192) [2.248]

divorced	−0.025	−0.278	0.362
	(0.209)	(0.259)	(0.358)
	[1.068]	[1.121]	[1.103]
widowed	0.143	−0.266	1.051*
	(0.278)	(0.335)	(0.501)
	[1.113]	[1.243]	[1.072]
number of children	0.012	−0.016	0.038
	(0.051)	(0.079)	(0.069)
	[2.280]	[2.540]	[2.245]
graduated from college	0.152*	0.137	0.157
	(0.084)	(0.121)	(0.117)
	[1.160]	[1.194]	[1.172]
income	0.132***	0.099***	0.154***
	(0.023)	(0.038)	(0.030)
	[3.062]	[2.545]	[2.039]
work hours (log)	−0.285***	−0.321**	−0.200
	(0.100)	(0.135)	(0.163)
	[1.721]	[1.719]	[1.497]
executive position	0.195	−0.336	0.278
	(0.224)	(0.532)	(0.249)
	[1.120]	[1.160]	[1.131]
temporary employed	0.020	−0.173	0.100
	(0.136)	(0.179)	(0.240)
	[2.396]	[2.543]	[1.665]
dispatched worker	−0.553*	−0.453	−1.013
	(0.300)	(0.343)	(0.646)
	[1.129]	[1.244]	[1.076]
self-employed	−0.020	−0.029	−0.021
	(0.162)	(0.307)	(0.193)
	[1.345]	[1.574]	[1.297]
family worker	0.085	−0.342	0.748**
	(0.253)	(0.330)	(0.346)
	[1.188]	[1.372]	[1.158]

employed at government agency	0.223	0.124	0.310
	(0.144)	(0.202)	(0.211)
	[1.129]	[1.154]	[1.142]
commuting time	−0.011	0.013	−0.017
	(0.019)	(0.032)	(0.024)
	[1.331]	[1.637]	[1.231]
satisfaction with health	0.148***	0.150**	0.136**
	(0.047)	(0.069)	(0.066)
	[1.320]	[1.253]	[1.366]
satisfaction with area of living	0.136***	0.227***	0.071
	(0.046)	(0.067)	(0.066)
	[1.278]	[1.310]	[1.343]
satisfaction with leisure	0.218***	0.146*	0.297***
	(0.050)	(0.075)	(0.068)
	[1.391]	[1.502]	[1.386]
member in civil/social organization	0.278*	0.278	0.291
	(0.150)	(0.226)	(0.202)
	]1.152]	[1.188]	[1.154]
member in hobby organization	0.168	0.159	0.218
	(0.128)	(0.176)	(0.191)
	[1.160]	[1.240]	[1.150]
frequency of meeting friends	0.018	0.013	0.016
	(0.038)	(0.059)	(0.050)
	[1.260]	[1.315]	[1.289]
frequency of doing sports	0.020	0.016	0.023
	(0.026)	(0.036)	(0.037)
	[1.141]	[1.112]	[1.236]
number of books/comics/magazines	0.120***	0.085	0.148***
	(0.036)	(0.053)	(0.051)
	[1.158]	[1.200]	[1.176]
hours watching TV	−0.013	−0.012	−0.013
	(0.026)	(0.037)	(0.039)
	[1.126]	[1.201]	[1.168]

constants:			
y>=1	5.460**	10.181***	−0.331
	(2.338)	(3.451)	(3.491)
y>=2	3.843*	8.615**	−2.014
	(2.333)	(3.442)	(3.483)
y>=3	1.913	6.635*	−3.925
	(2.332)	(3.436)	(3.482)
observations	2,209	1,064	1,145
McFadden R^2	0.049	0.039	0.063
χ^2	275.314***	105.493***	188.570***
	(df = 26)	(df = 25)	(df = 25)
log-likelihood	-2702.527	-1293.052	-1396.332
	(df=29)	(df=28)	(df=28)

4.2 Additional Considerations and Regression Models

4.2.1 Multicollinearity

Besides the variables *age* and *age*^{2},[39] the VIFs for the remaining variables are clearly under the critical values 5 or 10 in all estimated models. Therefore it can be assumed that the degree of multicollinearity is not seriously high.

4.2.2 Missing Values: Results from Proxy Variable Models

This subsection mainly describes how significance levels change, if the income variable is replaced with the proxy variable *satisfaction with one's financial situation.* The analysis is based on the regression tables 8 (OLS) and 9 (ordered logit), which can both be found in Section 3 of the appendix. These additional models were estimated, because they allow to analyze a bigger sample size. The reason for this is that many survey participants declined to give information on their salary, but agreed to report upon their satisfaction level with their financial situation. For instance, the general sample of the proxy models include over 400

[39]As explained in the Methods Chapter, this is not problematic, since it does not bias the signifcance levels.

additional observations.

Similar to the income variable in the original models, the proxy variable *satisfaction with one's financial situation* is positively associated with job satisfaction and highly significant.[40] The significance level and direction of the coefficients remain the same in both model types (original and proxy) for the variables *never married* and *number of books, comics or magazines.* In addition the variables *divorced, number of children, temporary employed, self-employed, family worker, commuting time, frequency of meeting friends, frequency of doing sports* and *hours of watching TV* remain insignificant in both model types. The significance level for the variable *widowed* remains mainly the same for both model types - except the male sample of the logit model, where it is higher for the proxy model.[41] Likewise, the significance and magnitude of the coefficients remains similar for the variables *dispatched worker*,[42] *satisfaction with one's health*,[43] *area of living*[44] and *leisure*.[45] Furthermore, the significance level and coefficient is lower for the variable *male* in the proxy models. However the variable is still significant at a 5% level and the direction of coefficient remains the same.

A bigger divergence in the significance levels between the original and proxy mod-

[40]The association with the proxy variable is significant at a 1% level for all samples and models. The income variable is significant at a 1% level for all models, except in the OLS model based on the female sample, where it is still highly significant at a 5% level.

[41]The variable is significant at a 10% level in the original model and significant at a 5% in the proxy version.

[42]In the original models the variable *dispatched worker* is mildly significant at a 10% level for the general sample in both the OLS and the logit model as well as for the male sample in the ordered logit model, while in the proxy models it was found to be significant only for the general sample (10% level in the OLS model and 5% level in the ordered logit model).

[43]In the original models the variable *satisfaction with one's health* is highly significant at a 1% level for the general sample as well as mostly significant at a 5% level for the remaining samples - only once (male sample, OLS model) the significance level falls to only 10%). In the proxy models the variable remains highly significant for the general and female sample, but turns to insignificant in the male sample models.

[44]The variable *satisfaction with one's living area* is highly significant at a 1% level for the general and female sample in the original models. In the proxy models the significance levels drop, but remain highly significant.

[45]In the original as well as in the proxy models *satisfaction with one's leisure* is significant at a 1% level for the general and male sample. For the female sample the variable is still significant in the original models (5% level in the OLS and 10% level in ordered logit model), but insignificant in the proxy models.

els were observed for the following variables: While being graduated from a college is only mildly significant at a 10% level for the general sample in the original models, the variable is highly significant at a 5% level in the proxy models based on the general as well as the male sample. Similarly, the variable *executive position*, which was identified even as insignificant for all samples in the original models, is significant at a 5% level for the general as well as the male sample of the proxy regression models. In addition, *employment at a government agency* and *membership in a civil or social organization* reach much higher significance levels in the proxy models.[46] In contrast, *work hours* are not significant in all proxy models, despite having been previously identified as highly significant for the general and female sample of the original models. Similarly the variables *age* and *age*^{2} are not significant anymore in the proxy variable models.

4.3 Irrelevant Variables: Results from Adjusted Models

To account for irrelevant variables the following method was used: Variables with the lowest significance were removed stepwise from several model until the adjusted R^2 started dropping. This procedure was applied separately on the different samples (general, female and male). Since this procedure can only be applied to OLS regression, only OLS models were estimated via this method. The final results are displayed in the tables 10, 11 and 12 in the appendix (Section C).

4.3.1 Results based on the Main Sample

The results for the main sample are displayed in table 10 of the appendix.

For the main general sample the variables *age, age*^{2}*, male, never married, income, work hours (log), number of books/comics/magazines* and *satisfaction with one's*

[46]In the original models, the variable *employed at government agency* is only mildly significant for the general sample in the ordered logit model. In the proxy models however, being employed at a government agency is positively associated with job satisfaction and highly significant for the general and male sample in both the OLS and the ordered logit models. Similarly, being a member in a social or civil organization is significant solely for the general sample in the original models (5% level in the OLS and 10% level in the ordered logit model), but highly significant for all general and male samples in the proxy models.

health, area of living or leisure were found to be significant at a 1% level. Furthermore the dummy variables *graduated from college, dispatched worker* and *member of civil/social organization* were identified as significant at a 5% level. Besides, the variables *employed at a government agency, membership in hobby organization* and *frequency of doing sports* were included in the model for the general sample, since the elimination of these variables would have led to a lower adjusted R^2.
For the female sample only the variables *never married, income* and *satisfaction with one's area of living* were detected as significant at a 1% level. *Work hours (log)* as well as the *satisfaction levels with health* as well as *leisure* were found to be significant at a 5% level, while *age, age*^{2} and *member of civil/social organization* were only significant at 10% level for the female sample. Aside from this, the variables *divorced, graduated from college, temporary, dispatched or family worker, commuting time, member in hobby organization* and *number of books/-comics/magazines* were incorporated in the final model.
The significance levels and selected variables in the model for the male sample differ considerably from the model based on the female sample. First, there are more variables, which were found to be significant at a 1% level (namely *never married, widowed, income, satisfaction with leisure* and *number of books/comics/magazines*). Also, partly other variables were identified as significant at a 5% level (*age*^{2}*, family worker* and *satisfaction with health*) and significant at a 10% level (*dispatched worker* and *member of civil or social organization*). Besides, *graduated from college, work hours (log), employed at government agency, satisfaction with one's area of living* and *member of hobby organization* were included in the model for the male sample.

4.3.2 Results based on the Subsamples A and B

Some variables - such as happiness level or solidarity at work place - are expected to be influential determinants of job satisfaction, but could not be included in the main models, since the corresponding questions only appeared on one of the two questionnaire forms. In order to account for these important variables separate models for the subsamples A and B were estimated. Again the method of stepwise deletion of irrelevant variables was used to increase the precision of the estimated coefficients for the explanatory variables. The tables 11 and 12 (Section C of the appendix) are structured in the same way as previous tables and display estimates separately for the general, female and male sample.

Considering the subsample A the following variables were included in all three models ("general", "female" and "male"): *income, executive position, self-reported happiness (A), anxiety regarding economic future* and *trust in people and organizations.* In addition, the variables *never married, work hours (log), family worker, satisfaction with area of living, member of civil/social organization, progression of economic situation in the last year* and the *believe in opportunities to improve one's life* were included in the "general" as well as the "female model". Among the variables which were included in the "male" and "general" model, but not in the "female" model, are: *widowed, satisfaction with leisure, frequency of meeting friends, number of books/comics/magazines, frequency of private trips, self-reported health* and *self-reported social status.* In contrast, the variables *gender* and *divorced* were only included in the "general" model and the variables *number of children, temporary employed* and *employed at government agancy* were only included in the "female" model.

Regarding the "general" model the variables *never married, income, work hours (log), progression of economic situation, self-reported status* and *anxiety regarding economic future* were identified as highly significant at a 1% level. Besides satisfaction with leisure, frequency of private trips and *self-reported happiness (A)* were found to be significant at a 5% level. Still significant at a 10% level are *satisfaction with area of living, member of civil/social organization* and *believe in opportunities to improve one's life.*

In respect to the "female" sample, the determinants *never married, self-reported happiness (A), believe in opportunities* and *anxiety regarding economic future* were

classified as significant at a 1% level. *Work hours (log), family worker* and *satisfaction with area of living* were determined as significant at a 5%. *Membership in civil/social organizations* was the only variable in the "female" model which was significant at a 10% level.
For the "male" model *income, satisfaction with leisure* and *self-reported social status* fall into the group of variables highly significant at a 1% level. Furthermore the variables *number of books/comics/magazines, frequency of private trips, anxiety regarding economic future* as well as *trust in people and organizations* were identified as significant at a 5% level and *widowed* as well as *self-reported health* as significant at a 10% level.

Relating to the subsample B the variables *age*^{2}*, self-reported happiness (B), solidarity at work place, progression of economic situation* and *believe in opportunities to improve one's life* were incorporated in the general and both gender-specific models. Moreover, *college graduation, commuting time* and *loneliness* were included in the "female" as well as the "general" model. Furthermore, in the "male" as well as the "general" model the variables *never married, income, executive position, dispatched, family worker, satisfaction with leisure, number of books/-comic/magazines, pain interfering with daily activities* and *anxiety regarding one's economic future* were used. In addition, *self-employment* was included in the "female" as well as the "male" model, but not in the "general" model. In contrast, *widowed, member of civil/social organization* and *self-reported social status* were only included in the "male" model; *temporary employment, satisfaction with health* and *frequency of meeting friends* only in the "female" model; and *work hours (log)* only in the "general" model.
Considering the general sample the variables *age*^{2}*, male, income, solidarity at work place, progression of economic situation* and *believe in opportunities to improve one's life* were found to be significant at a 1% level. Significant at a 5% level were the *number of books/comics/magazines* (read per week), *pain interfer-*

ing with daily activities and *self-reported happiness.* Besides, *college graduation* and *anxiety regarding economic future* were mildly significant at a 10% level. Regarding the female sample *age*^{2}*, loneliness* and *solidarity at work place* were significant at a 1% level and *college graduation* as well as *commuting time* were significant at a 5% level. Still significant at a 10% level were *self-employment* and *progression of one's economic situation in the last years.*

In respect to the male subsample *age*^{2}*, solidarity at work place* and *progression of one's economic situation in the last years* were identified as highly significant at a 1% level. Additionally, *income, dispatched, number of books/comics/magazines, pain interfering with daily activities* and *believe in opportunities to improve one's life* were classified as significant at a 5% level and *widowed, executive position* and *family worker* as significant at a 10%.

In the next chapter the combined results of all estimated models are explained. Furthermore they are compared to the study findings previously described in the literature review. Furthermore the limitations of this study and recommendations for future research are outlined.

5 Discussion

Job satisfaction is - besides its normative value - of considerable economic importance. However, there is surprisingly little literature on its determinants - especially in respect to non-western countries.

One of this study's main purposes is to provide a detailed overview of known demographic, economic and psychological (e.g. measures of subjective well-being) determinants and their possible effects on job satisfaction. For this purpose micro-data from the General Japanese Social Survey of the year 2010 is used as a data source. Since the vast majority of comparable studies has been conducted based on survey data from the US or the EU, the overview by itself is a valuable addition to the academic literature. However, the study provides a number of other scientific contributions. One of these contributions is to look at relationships between job satisfaction and control variables which have not been explored so far by other studies - for reasons such as a lack of data. Examples for such control variables are *satisfaction level with one's leisure activities* and different *types of leisure activities*. In addition, this study is designed to find potential gender differences regarding the effects of the selected variables. Finally, on the meta-level the thesis provides data on comparability of the results from ordered logit regression and OLS regression.

To perform these multiple tasks - and to account for several statistical issues explained in section 3.6 - several regression analyses, which access the relationship between job satisfaction and different sets of control variables, were conducted. In total, 21 regression models were estimated.

5.1 Combined Results of all Models

This section summarizes the results of all for this study estimated models.

The findings are consistent with the literature review for the following variables:

Age: Considering the main samples, *age* as well as age^2 are significant for most models, indicating a U-shaped relationship. Furthermore, the significance levels are stronger for the ordered logit model. However, this relationship disappears when either a self-reported happiness variable is included in the model or the income variable is replaced by the proxy variable *satisfaction with one's financial situation* - which is a subjective wellbeing variable itself and thus likely to be highly correlated with self reported happiness. This observation reveals that the U-shaped relationship, which has been found by many studies beforehand, might not be direct, but result from omitting the self-reported happiness variable. This hypothesis is plausible, since a U-shaped relationship between self-reported happiness and age, is a highly reoccurring pattern in studies on general happiness. Unfortunately, the subsample models, which include a self-reported happiness variable, do not give a conclusive picture of the real nature behind the relationship of age and job satisfaction: On the one hand, the models based on subsample A do not identify the age variable as significant enough to be included in the model. On the other hand, in the models based on subsample B age^2 turns out to be highly significant at a 1% level.

Gender: In all models - except the one based on subsample B - the gender variable was highly significant with the male participants being on average less satisfied than their female counterparts. A negative relationship between being male and job satisfaction was also reported by most studies cited in the literature review.

Number of children was not significant in any model of this study. No significant relationship was also found by most other studies mentioned in the review.

Income and *satisfaction with financial situation* are both positively correlated with job satisfaction and highly significant in almost every model. However, the coefficients indicate that job satisfaction depends to a much higher degree on income for male workers than for their female counterparts. For *work hours (log)* it is exactly the opposite. The variable is highly significant for the majority of general and female samples, but never for the male samples. In the models where the variable was identified as significant, it is negatively related to job satisfaction. A big body of literature confirms the negative relationship between work hours (log) and job satisfaction, but only one of the reviewed studies finds the relationship to be stronger for female workers. In contrast, both - positive relationship and gender difference - is confirmed for the income variable by a number of previous studies.

Commuting time: was found to be significant in only one model (based on the female subsample B). This is not enough evidence to support the thesis of a negative effect of commuting on job satisfaction. Since, there are very few studies which have explicitly looked at the relation between commuting time and job satisfaction, this finding is not in contradiction with the literature review.

Health: There are three different health variables, which were included in some of the models;[47] Namely *satisfaction with one's health condition, self-reported health* and *pain interfering with daily activities.* Since the health satisfaction variable was measured for the whole sample, it was included in most models, while the other two measures were only included in very few models basing on either subsample A or B. The health satisfaction variable was positively associated with job satisfaction and highly significant in 81.25% of the cases, in which it was included in the model. In the proxy models the variable's significance level dropped considerably, especially for the male samples. The health satisfaction variable was only included in one of the models based on the subsamples,[48] since the alternative health variables *self-reported happiness* and *pain interfering with daily activities* were identified as more significant in these kind of models. An interesting observation is that the chosen health variables remain mostly insignificant in the models based on the subsample A, but are found to be highly significant in the models based on subsample B.[49] The evidence for positive relationship between health and job satisfaction, is one of the strongest in this thesis, since in approximately 71% of the estimated models a health variable has been determined as highly

[47]Per model only one health variable was included.

[48]See subsample B, female sample.

[49]Unfortunately, the reason for this could not be determined. Even after a test in which previously possibly omitted variables were removed from a model both possible health variables remained insignificant.

significant. These findings are consistent with the academic literature, which confirms a positive relationship between health and job satisfaction.

Satisfaction with area of living is positively related to job satisfaction and was identified as highly significant in 6 out of 7 of estimated models based on the female samples and in 5 "general models", but never for a model based on the male sample in this study. Although previous studies have also found a positive relationship between housing situation and job satisfaction, this study is - as far as the researcher knows - the first to identify a gender difference of the effect.

Self-reported happiness (A as well as B) is positively related to job satisfaction and is highly significant for both general samples. The positive relationship is supported by previous studies, which suggest a positive reciprocal relationship between both variables. However, while the self-reported happiness variable in this study is also significant at a 1% level for the female and at a 10% level for the male sample in the models based on subsample A, *self-reported happiness* remains insignificant for the gender-specific models based on subsample B. This raises the question by how much the estimates of a control variables can be influenced through the wording of the question and in which way?

Partly consistent with previous findings are the results on the following variables:

Marital status: In all models, participants, who have never been married tend to be less satisfied with their job than survey participants who are married, divorced or widowed. For most models, this relationship is highly significant and tend to be much stronger for the female samples. Interestingly, widowed men and married women are the groups who were found to be most satisfied with their job. These results are just partly consistent with most studies that were reviewed, since it is only the *never married* group, that is significantly less satisfied with their jobs in this study, and not those who are divorced or widowed like in most previous studies.

The results are in contradiction with the majority of previous research findings in respect to the following variables:
College graduation was found to be positively correlated with job satisfaction and is significant for the majority of the models of this thesis. However, a positive relationship have only be found in one of the reviewed studies.
Furthermore, the *type of employment* did not matter to such a high degree as expected. However, some patterns, especially considering the different gender groups, were interesting to observe. For example, though being in an *executive position* is never significant at a high level as long as the income variable is included in the model, one pattern appears in most models; Men in an executive position ceteris paribus, tend to be more satisfied with their job, while women tend to be less satisfied than their counterparts in regular positions. This tendency is even stronger for the family worker variable, which is highly significant for the majority of models based on a male sample. The only two variables which are significant at an at least 10% level for the majority of the general samples are *dispatched* and *employment at government agency*, since their coefficients go in the same direction for either gender. However, the association with job satisfaction - negative for *dispatched* and positive for *employment at government agency* - are again more pronounced in respect to the male samples. The findings are surprising, since the reviewed research did not include these two variables (unless *dispatched* can be counted as a form of temporary work). Furthermore, variables which were suggested to be important in previous studies (e.g. executive position, self-employment and temporary work) turned out to be insignificant for this study.

Considering the variables which were specifically investigated for this thesis the following relationships were identified:
Progression of economic situation was classified as positively associated with job satisfaction and highly significant for both general samples, but only for the gender-specific samples from subsample B. Moreover, *believe in opportunities to*

improve one's life was positively correlated with job satisfaction and highly significant for half of the subsample models. Furthermore, a highly significant negative association was found between *anxiety regarding one's economic future* and job satisfaction for the models based on subsample A. Unfortunately, *solidarity at work place* was only assessed via questionnaire form B, which was why it could only be included in the corresponding models. However, in these models its positive correlation with job satisfaction was not only highly significant at a 1% level, but its p-value was also the smallest among all included variables, making the variable very interesting for future research. In regard to the leisure related variables the following relationships were observed:

Satisfaction with leisure was included in 19 out of the 21 estimated models. Its estimates have usually positive values and in approximately 86% of all "male" and "general models" the variable was classified as highly significant (usually at a 1% level). Considering the female samples, the variable was found to be significant at an at least 10% level in 4 out of 7 models. If *self-reported happiness (B)* and *pain interfering with daily activities* are included in a model, the variable's significance level strongly decreases, which is why *satisfaction with leisure* is not significant in the models based on subsample B.

In regard to the considered leisure activities, *frequency of meeting friends, doing sports, spending time in nature* and *watching TV* remained insignificant for all models. In contrast, *number of books/comics/magazines* was registered as highly significant for all models based on a male sample, and for all but one model based on a general sample. *Frequency of private trips*, which was only included in the models based on subsample A, was found to be significant at a 5% level for the male and general sample. Finally, *membership in a civil or social organization* was positively associated with job satisfaction and highly significant for the majority of the models based on the general sample. This specific finding is not so surprising, since previous studies have suggested that people who do voluntary work report higher life satisfaction (Argyle, 1999, p. 365). However, there is also evidence that happier people are more willing to contribute to other people's well-being (Frey, 2008, p. 24). This means that - like for most variables in this study - the determination of the causal direction of the effect is tricky, since there are plausible arguments for both directions or a reciprocal relationship.

5.2 Limitations and Recommendations for Future Research

As indicated above, the main limitation of this study is that due to the cross-sectional character of the data and difficulty to find good instruments for the exogenous variables, it is not possible to determine the direction of the causal relation between the independent and dependent variables. Despite this constraint, the findings of this thesis are still quite meaningful, because the study makes multiple and unique contributions to the academic literature - e.g. on the comparability of OLS and ordered logit models and gender differences. In particular, the study sheds light on rarely investigated determinants of job satisfaction (e.g. leisure satisfaction, types of leisure activities, solidarity at work place, etc.,). However, since the data is derived from a Japanese survey, the results considering these variables cannot be transferred to other cultural contexts such as developing or western countries.

Based on these shortcomings recommendations for future research can be derived:
1. To account for the endogeneity problem experimental or panel data need to be used. For example, to investigate the effect of leisure activities on job satisfaction an experimental study could be designed, in which a group of workers is instructed to engage in a specific leisure activity or spend more time on it (e.g. reading, volunteer work, etc.,). Then the results could be compared to a control group in order to estimate the magnitude of the causal effect on job satisfaction. However, a study like this might be difficult to perform, since it would constitute a challenging task to incentivize the participants to change their behavior in their leisure time. Therefore, the use of panel data might be an easier approach to account for endogeneity in future studies on job satisfaction.

2. To determine, whether the unique findings of this study also apply to other countries, a similarly designed regression analysis needs to be run on international data. In particular, the following questions seem to be worth cross checking with international data:

a. Is the U-shaped relationship between age and job satisfaction direct or does it always disappear when a self-reported happiness variable is included in the model?

b. Is the negative effect of work hours (log) always stronger for women and the effect of income stronger for men?

c. Is the correlation between *solidarity at work place* and job satisfaction always so strong and significant like in the models of this study, or is this a phenomenon which is more prevalent in collectivistic countries like Japan?

d. Is it a coincidence that *number of books, magazines or comics* and *membership in civil or social organization* are the leisure activities with the highest significance levels, or are these findings reproducible with other survey data. If yes, what might be the reason for this?

The researcher hopes that this thesis inspires more interest in the research on the determinants of job satisfaction and that current as well as future findings in this field may contribute to better working environments and productivity in companies around the globe.

Appendices

A Definitions of all Variables

Variable	*Definition*
job satisfaction	Ordinal variable accessing how satisfied the respondents are with their with their job. In the survey the variable is assessed via question Q13 of the interview (“on the whole, how satisfied are you with the (main) job you have?/genzai no shigoto ni dono kurai manzoku shite imasu ka”), which requests the participants to give an integer rating on a 5-point-scale. Missing values (no answer, not applicable and don’t know) were deleted. Furthermore, the original categories “dissatisfied” and “somewhat dissatisfied” were combined to one category. In addition, the original order was reversed, so that a higher numerical value corresponds to a higher level of satisfaction. The four categories are: 1. Dissatisfied and somewhat dissatisfied/ fuman de aru + dochira ka to ieba fuman de aru 2. Neither satisfied nor dissatisfied/dochira tomo ienai 3. Somewhat satisfied/dochira ka to ieba manzoku shite iru 4. Satisfied/manzoku shite iru
age	Respondent’s age in years (20-69). Respondents over the age of 69 were removed from the sample due to the lack of observations.
male	Dummy variable with a value of 1, if the respondent is male.

marital status:	Missing values were deleted. Furthermore, those who answered with "separated" were excluded due to a lack of observations. The reference value is married.
never married	Dummy variable with a value of 1, if the respondent was never married.
divorced	Dummy variable with a value of 1, if the respondent is divorced.
widowed	Dummy variable with a value of 1, if the respondent is widowed.
number of children	Missing values were deleted. Ordinal variable with 5 possible answers: no, one, two, three and four or more children.
graduated from college	Dummy variable, with the value one if the respondent graduated from college. Missing values were deleted. Besides, "higher school or vocational school in the old system," "normal school in the old system," "vocational school/commerce school in the old system" were excluded due to a lack of comparability and observations.
income	Respondent's yearly income. Originally 19 categories: 1. None 2. Less than 700,000¥ 3. 700,000¥ - 1 million¥ 4. 1 million¥ - 1.3 million¥ 5. 1.3 million¥ - 1.5 million¥ 6. 1.5 million¥ - 2.5 million¥ 7. 2.5 million¥ - 3.5 million¥ 8. 3.5 million¥ - 4.5 million¥ 9. 4.5 million¥ - 5.5 million¥ 10. 5.5 million¥ - 6.5 million¥ 11. 6.5 million¥ - 7.5 million¥ 12. 7.5 million¥ - 8.5 million¥ 13. 8.5 million¥ - 10 million ¥ 14. 10 million¥ - 12 million¥ 15. 12 million¥ - 14 million¥

16. 14 million¥ - 16 million¥
17. 16 million¥ - 18.5 million¥
18. 18.5 million¥ - 23 million¥
19. 23 million¥ or over

For the analysis, missing values (don't want to state the income, don't know, no answer and no income) were removed. Furthermore values equal to 14 or higher were combined to a single category (equal to or over 10 million¥) due to the limited number of observations.

work hours (log)	Natural log of the continuous variable weekly work hours ranging from 1 to 112. Missing values (not applicable and no answer) were deleted.
type of job	Missing values (don't know, not applicable and no answer) were deleted. Furthermore unemployed and retired participants, students as well as those who were exclusively engaged in home keeping or not able to work due to physical and health problems were removed from the sample. The reference value is "regularly employed/jōji koyō no ippan jūgyōsha."
executive position	Dummy variable with a value of 1, if the respondent is an executive of a company or a corporation (jap: keieisha + yakuin).
temporary employed	Dummy variable with a value of 1, if the respondent is a temporary worker including daily workers and part-time temporary workers, doing piece work at home (jap: rinji koyō, pāto arubaito, naishoku).
dispatched worker	Dummy variable with a value of 1, if the respondent is a dispatched worker from a temporary personnel agency (jap: haken shain).
self-employed	Dummy variable with a value of 1, if the respondent is self-employed (jap: jieigyō-nushi + jiyū gyōsha).
family worker	Dummy variable with a value of 1, if the respondent is a family worker (jap: kazoku juūgyō-sha).
employed at government	Dummy variable with a value of 1, if the respondent

agency	is working at a government agency (jap: kankōchō).
commuting time	Missing values (not applicable and no answer) were deleted. The continuous variable were coded into an ordinal variable with ten categories: 1. Less than 9 minutes 2. 10-19 minutes 3. 20-29 minutes 4. 30-39 minutes 5. 40-49 minutes 6. 50-59 minutes 7. 60-69 minutes 8. 70-79 minutes 9. 80-89 minutes 10. at least 90 minutes.
satisfaction with health	Ordinal variable accessing how satisfied the respondents are with their health and physical condition. The original question (Q12F on the questionnaire form A and B) to this variable was stated as "how much satisfaction do you get from the following areas of life? Your health and physical condition:/ Seikatsu men ni kan suru ika no kōmoku ni tsuite, anata wa dono kurai manzoku shi te imasu ka. Kenkō jōtai." The respondents were instructed to give a rating on a 5-point-integer scale. However, for the models the original order was reversed, so that a higher numerical value corresponds to a higher level of satisfaction. This means the lowest possible value "0" equates "dissatisfied/fuman" and the highest value "4" equates "satisfied/manzoku."
satisfaction with area of living	Ordinal variable accessing how satisfied the respondents are with their living place (jap: sunde iru chiki). The respondents were instructed to give a rating on a 5-point-integer scale. Again the original order was reversed, so that a higher numerical value corresponds to a higher level of satisfaction. This means the lowest possible value "0" equates "dissatisfied/fuman" and the highest value "4" equates "satisfied/manzoku."

satisfaction with leisure	Ordinal variable accessing how satisfied the respondents are with their non-work activities (jap: yoka no sugoshikata). The respondents were instructed to give a rating on a 5-point-integer scale. The original order was reversed for the models, so that a higher numerical value corresponds to a higher level of satisfaction. This means the lowest possible value "0" equates "dissatisfied/fuman" and the highest value "4" equates "satisfied/manzoku."
member of civil or social organization	Dummy variable with a value of 1, if the respondent is a member in a citizens' movement (jap: shimin undō), a consumers' cooperative (jap: shōhisha undō no gurūpu) or a social service group (jap: borantia no gurūpu). Missing values were deleted.
member in hobby organization	Dummy variable with a value of 1, if the respondent is a member in a hobby organization (e.g. chorus, photography, mountain hiking, etc) - jap: shumi no kai (korāsu, shashin, yamāruki nado). Missing values were deleted.
frequency of meeting friends	Ordinal variable accessing how often the respondents meet their friends. Missing values were removed. The original order was reversed for the models, so that a higher numerical value corresponds to a higher frequency of meeting friends. Furthermore, the original categories "almost everyday" and "several times a week" were combined to one category. The six categories are: 1. Never/mattaku shite inai 2. About once a year/toshi ni ikkkai teido 3. Several times a year/toshi ni su kai 4. About once a month/tsuki ni ikkai teido 5. About once a week/shu ni ikkai teido 6. Several times a week (combined)/shū ni sūkai.

frequency of doing sports	Ordinal variable accessing how often the respondents engage in exercise or playing sports. Missing values were removed. Furthermore, the original order was reversed for the models, so that a higher numerical value corresponds to a higher frequency of doing. The five categories are: 1. Scarcely any exercise/hotondo shinai 2. Several times a year/toshi ni sūkai-teido 3. About once a month/tsuki ni ikkai teido 4. About once a week/shū ni ikkai teido 5. More than several times a week/shū ni sūkai ijō.
number of book, comics and magazines	Ordinal variable accessing the number of books, comics or magazines which the respondent reads per month. Missing values were deleted. The lowest observed value is "0" and the highest observed value is "4."
hours watching TV	Ordinal variable measuring the amount of time the respondent spends watching TV per day. Missing values were deleted. There are eight categories with "zero hours" as the lowest and "seven hours and more" as the highest.
satisfaction with financial situation	Ordinal variable accessing how satisfied the respondents are with the current financial situation of their household. The original question (Q12D on the questionnaire form A and B) to this variable was stated as "how much satisfaction do you get from the following areas of life? D: The current financial situation of your household:/seikatsu men ni kan suru ika no kōmoku ni tsuite, anata wa dono kurai manzoku shi te imasu ka. D: Genzai no kakei jōtai." The respondents were instructed to give a rating on a 5-point-integer scale. However, for the models the original order was reversed, so that a higher numerical value corresponds to a higher level of satisfaction. This means the lowest possible value "0" equates "dissatisfied/fuman" and the highest value "4" equates "satisfied/manzoku."

frequency of private trips

Ordinal variable accessing how often in a year respondents go on trips of more than two days (business trips are excluded). It bases on Q6 of questionnaire form A. The Japanese phrasing of the question is: "Anata wa, donogurai no hindo de ichi tomari ijou no ryoku (shigoto igai) wo shimasu ka." Missing values (no answer and not applicable) were deleted and the original order was reversed for the models, so that a higher numerical value corresponds to more trips. Besides, the original categories "several times a month" and "at least once in month" were combined into one category. The four categories for the models are:

1. Never/mattaku shinai
2. About once a year/toshi ni ikkai teido
3. Several times a year/toshi ni sūkai
4. At least once in a month/tsuki ni ikkai teido.

self-reported health

Ordinal variable measuring how the respondents rate their health condition. The original question (Q8 on questionnaire form A) to this variable was stated as "how would you rate your health condition?/anata no genzai no kenkō jōtai wa, ikagadesu ka." The respondents were instructed to give a rating on a 5-point-integer scale. For the models the original order was reversed, so that a higher numerical value corresponds to a to a better health rating. Furthermore the two lowest health categories were combined into one category due to the lack of observations. This means the lowest possible value "0" equates "poor/warui" and the highest value "3" equates "good/yoi."

self-reported happiness (A)	Ordinal variable accessing how happy the respondents feel. The original question (Q13 on questionnaire form A) to this variable was stated as "are you happy?/anata wa, genzai shiawase desu ka." The respondents were instructed to give a rating on a 5-point-integer scale. The original order was reversed for the models, so that a higher numerical value corresponds to a higher happiness level. This means the lowest possible value "0" equates "unhappy/fushiawase" and the highest value "4" equates "happy/shiawase."
progression of economic situation in last years	Ordinal variable accessing how the financial situation of the respondents has developed during the last years. The corresponding question is: "During the last few years, has your financial situation been getting better, worse, or has it stayed the same?/ kono 2-3 nen no aida ni, anata no keizaijoutai ha dou kawarimashita ka." It is stated as Q27 on questionaire form A and as Q17 on questionaire form B. Missing values were deleted and the original order was modified for the models, so that the highest numerical value corresponds with an improved financial situation. The three categories for the models are: 1. situation got worse/waruku natta 2. situation stayed the same/kawaranai 3. situation got better/yoku natta.
believe in opportunities to improve one's life	Ordinal variable accessing if the respondents believe that the Japanese society offers enough opportunities for them to increase their and/or their family's standard of living (Japanese question: Ima no nihon no shakai ni wa, anata ya anata no kazoku no seikatsu suijun o kōjō sa seru kikai ga, donokurai arimasu ka). The variable was derived from question Q31on form A and Q20 on form B. The respondents were instructed to give a rating on a 5-point-integer scale. Missing values were deleted and the original order was reversed for the models, so that a higher numerical value corresponds to a stronger believe in one's opportunities. The five categories for the models are:

1. Not sufficient at all/mattaku nai
2. Not very sufficient/amari nai
3. Neither nor/dochira tomo ienai
4. Somewhat sufficient/sukoshi wa aru
5. Sufficient/jūbun ni aru.

self-reported social status	Ordinal variable accessing to which social strata respondents count themselves. The wording of the question (Q33 on form A and Q21 on form B) from which this variable was derived is: "If we were to divide the contemporary Japanese society into the following five strata, which would you say you belong to?/kari ni genzai no nihon no shakai zentai wo, ika no itsustsu no sou ni wakeru to sureba, anata jishin ha, dore ni hairu to omoimasu ka." The respondents were instructed to pick a category among five. The original order of theses categories was reversed for the models, so that a higher numerical value corresponds to a higher social strata. The five categories for the models are: 1. Lower/shita 2. Lower middle/chū no shita 3. Middle/chū no naka 4. Upper middle/chū no ue 3. Upper/ue.
anxiety regarding economic future	Ordinal variable accessing how anxious respondents feel about their economic situation in the future. The corresponding question (Q34 on form A and Q22 on form B) was posed as follows: "Do you feel anxious about your economic situation in the future?/kongo no seikatsu ni tsuite, keizaiteki ni fuan wo kanjite imasu ka." The respondents were instructed to give a rating on a 5-point-integer scale about how anxious they feel. Missing values were deleted and the original order was reversed for the models, so that higher numerical values corresponds with a higher anxiety. The five categories for the models are: 1. I don't feel anxious at all/mattaku kanjite inai 2. I don't feel anxious very much/amari kanjite inai 3. I have mixed feelings/dochira-tomo ienai

4. I feel somewhat anxious/aruteido kanjite iru
5. I feel very anxious/totemo kanjite iru.

trust in people and organizations	Ordinal variable accessing how much respondents trust in other people.The belonging question (Q67 on form A) is stated as follows: "generally speaking, would you say that most people can be trusted?/ippanteki ni, hito ha shinyō dekiru to omoimasuka". The respondents were given three options, which were brought in an ordinal order for the models of the thesis. Missing values were deleted. The three categories for the models are: 1. No/iie and yes/hai 2. Depends/baai ni yoru 3. yes/hai.
loneliness	Ordinal variable accessing how lonely respondents feel. The variable was derived from question Q66C on the questionnaire form B and was stated as follows: "How much do you feel the followings in your daily life? Loneliness/higoro no seikatsu de, anata wa ika no koto o, dono kurai kanjite imasu ka: kodokukan." Missing values were deleted and the original order was reversed for the models, so that higher numerical values correspond to a stronger feeling of loneliness. The four categories for the models are: 1. Not at all/mattaku kanjitte inai 2. To some extend/aru teido kanjitte iru 3. Not so much/amari kanjitte inai 4. Very much/totemo kanjitte iru.
pain interfering with daily activities	Ordinal variable accessing whether and how much pain interferes with the respondents' work activities. The original question (Q38 on questionaire form B) to this variable was stated as "during the past 4 weeks, how much did pain interfere with your normal work (including both work outside the home and housework)?/kako ikka getsu ni, itsumo no shigoto (kaji mo fukumimasu) ga itami no tame ni, donokurai samatageraremashita ka." The four categories are:

	1. Not at all/zenzen samatagerarenakatta 2. A little bit/wazuka ni samatage rareta 3. Moderately/sukoshi samatage rareta 4. Quite a bit/kanari samatage rareta 5. Extremely/hijou ni samatagerareta.
self-reported happiness (B)	Ordinal variable accessing how happy the respondents feel in general. The original question (Q13) to this variable was stated on the questionnaire form B and reads as "in general, are you happy?/zentai toshite, anata ha, genzai shiawase desu ka." The respondents were instructed to give a rating on a 5-point-integer scale. The original order was reversed for the models, so that a higher numerical value corresponds to a higher happiness level. This means the lowest possible value "1" equates "very unhappy/hijou ni fushiawase" and the highest value "5" equates "very happy/hijou ni shiawase."
solidarity at work place	Ordinal variable accessing the quality of the relationship between the respondents and their coworkers. The variable is based on Q33A on the questionnaire form B ("people in my workplace have a strong sense of solidarity/watakushi no shokuba deha, hitobito no aida no rentaikan ga tsuyoi"). Respondents could agree with the statement on a four point scale. Missing values (not employed, no co-workers, not applicable and no answer) were deleted and the original order was reversed for the models, so that higher numerical values correspond with a stronger solidarity. The four categories for the models are: 1. Untrue/sou omowanai 2. Somewhat untrue /dochira ka to ieba sou omowanai 3. Somewhat true/dochira ka to ieba sou omou 4. True/sou omou.

B Descriptive Statistics

Table 3: Means and SEs of Variables (Main Sample)

		n	mean	SE of mean	min.	1nd qu.	median	3rd qu.	max.
job satisfaction	general	2209	1.817	0.020	0	1	2	3	3
	women	1064	1.865	0.028	0	1	2	3	3
	men	1145	1.773	0.028	0	1	2	2	3
age	general	2209	45.65	0.268	20	36	46	56	69
	women	1064	44.64	0.382	20	35	45	55	69
	men	1145	46.58	0.373	20	37	47	57	69
male	general	2209	0.518	0.011	0	0	1	1	1
	women	1064	0	0	0	0	0	0	0
	men	1145	1	0	1	1	1	1	1
never married	general	2209	0.197	0.009	0	0	0	0	1
	women	1064	0.202	0.012	0	0	0	0	1
	men	1145	0.192	0.012	0	0	0	0	1
divorced	general	2209	0.054	0.005	0	0	0	0	1
	women	1064	0.077	0.008	0	0	0	0	1
	men	1145	0.032	0.005	0	0	0	0	1
widowed	general	2209	0.022	0.003	0	0	0	0	1
	women	1064	0.034	0.006	0	0	0	0	1
	men	1145	0.01	0.003	0	0	0	0	1

number of children	general	2209	1.545	0.024	0	0	2	2	≥4
	women	1064	1.546	0.034	0	0	2	2	≥4
	men	1145	1.544	0.033	0	0	2	2	≥4
graduated from college	general	2209	0.426	0.011	0	0	0	1	1
	women	1064	0.416	0.015	0	0	0	1	1
	men	1145	0.436	0.015	0	0	0	1	1
income	general	2209	6.805	0.067	≤2	4	7	9	≥14
(for definitions of categories	women	1064	5.025	0.076	≤2	3	5	7	≥14
see Section A of appendix)	men	1145	8.459	0.083	≤2	7	8	10	≥14
work hours	general	2209	39.49	0.342	1	30	40	49.5	112
	women	1064	32.66	0.434	1	21	35	40	92
	men	1145	45.85	0.446	3	40	45	50	112
executive position	general	2209	0.052	0.005	0	0	0	0	1
	women	1064	0.022	0.005	0	0	0	0	1
	men	1145	0.080	0.008	0	0	0	0	1
temporary employed	general	2209	0.269	0.009	0	0	0	1	1
	women	1064	0.456	0.015	0	0	0	1	1
	men	1145	0.095	0.009	0	0	0	0	1
dispatched worker	general	2209	0.020	0.003	0	0	0	0	1
	women	1064	0.031	0.005	0	0	0	0	1
	men	1145	0.010	0.003	0	0	0	0	1
self-employed	general	2209	0.083	0.006	0	0	0	0	1
	women	1064	0.055	0.007	0	0	0	0	1

	men	1145	0.109	0.009	0	0	0	0	1
family worker	general	2209	0.034	0.004	0	0	0	0	1
	women	1064	0.055	0.007	0	0	0	0	1
	men	1145	0.015	0.004	0	0	0	0	1
employed at government agency	general	2209	0.090	0.006	0	0	0	0	1
	women	1064	0.093	0.009	0	0	0	0	1
	men	1145	0.087	0.008	0	0	0	0	1
commuting time	general	2209	2.434	0.050	0	1	2	3	9
(for definitions of categories	women	1064	2.017	0.062	0	1	1	3	9
see Section A of appendix)	men	1145	2.821	0.076	0	1	2	4	9
satisfaction with health	general	2209	2.506	0.021	0	2	2	3	4
	women	1064	2.569	0.031	0	2	3	3	4
	men	1145	2.448	0.030	0	2	2	3	4
satisfaction with area of living	general	2209	2.753	0.022	0	2	3	4	4
	women	1064	2.722	0.032	0	2	3	4	4
	men	1145	2.783	0.030	0	2	3	4	4
satisfaction with leisure	general	2209	2.406	0.022	0	2	2	3	4
	women	1064	2.399	0.031	0	2	2	3	4
	men	1145	2.411	0.030	0	2	2	3	4
member in civil/social organisation	general	2209	0.076	0.006	0	0	0	0	1
	women	1064	0.062	0.007	0	0	0	0	1
	men	1145	0.088	0.008	0	0	0	0	1
member in hobby organization	general	2209	0.111	0.007	0	0	0	0	1

	women	1064	0.121	0.010	0	0	0	0	1
	men	1145	0.102	0.009	0	0	0	0	1
frequency of meeting friends	general	2209	2.536	0.025	0	2	3	3	5
(for definitions of categories	women	1064	2.648	0.035	0	2	3	3	5
see Section A of appendix)	men	1145	2.431	0.036	0	2	2	3	5
frequency of doing sports	general	2209	1.436	0.034	0	0	1	3	4
(for definitions of categories	women	1064	1.278	0.049	0	0	0	3	4
see Section A of appendix)	men	1145	1.583	0.048	0	0	1	3	4
number of books/comics/magazines	general	2209	0.887	0.025	0	0	0	1	4
	women	1064	0.902	0.036	0	0	1	1	4
	men	1145	0.873	0.035	0	0	0	1	4
hours watching TV	general	2209	2.923	0.033	0	2	3	4	$\geq$7
	women	1064	3.060	0.051	0	2	3	4	$\geq$7
	men	1145	2.797	0.044	0	2	3	4	$\geq$7

Table 4: Means and SEs of Variables (models using income proxy variable)

		n	mean	SE of mean	min.	1nd qu.	median	3rd qu.	max.
job satisfaction	general	2697	1.819	0.018	0	1	2	3	3
	women	1302	1.852	0.025	0	1	2	3	3
	men	1384	1.787	0.025	0	1	2	2	3
age	general	2697	45.76	0.242	20	36	46	56	69
	women	1302	45.07	0.347	20	35	46	55	69
	men	1384	46.45	0.339	20	37	47	57	69
male	general	2697	0.485	0.010	0	0	0	1	1
	women	1302	0	0	0	0	0	0	0
	men	1384	1	0	1	1	1	1	1
never married	general	2697	0.198	0.008	0	0	0	0	1
	women	1302	0.196	0.011	0	0	0	0	1
	men	1384	0.198	0.011	0	0	0	0	1
divorced	general	2697	0.052	0.005	0	0	0	0	1
	women	1302	0.073	0.007	0	0	0	0	1
	men	1384	0.032	0.005	0	0	0	0	1
widowed	general	2697	0.022	0.003	0	0	0	0	1
	women	1302	0.036	0.005	0	0	0	0	1
	men	1384	0.009	0.003	0	0	0	0	1
number of children	general	2697	1.548	0.022	0	0	2	2	≥4
	women	1302	1.571	0.031	0	0	2	2	≥4

	men	1384	1.531	0.030	0	0	2	2	$\geq$4
graduated from college	general	2697	0.428	0.010	0	0	0	1	1
	women	1302	0.416	0.014	0	0	0	1	1
	men	1384	0.436	0.013	0	0	0	1	1
satifaction with financial situation	general	2697	1.994	0.022	0	1	2	3	4
(for definitions of categories	women	1302	1.992	0.032	0	1	2	3	4
see Section A of appendix)	men	1384	1.996	0.030	0	1	2	3	4
work hours	general	2697	39.41	0.306	1	30	40	49	112
	women	1302	32.67	0.395	1	22	35	40	100
	men	1384	45.72	0.397	2	40	45	50	112
executive position	general	2697	0.053	0.004	0	0	0	0	1
	women	1302	0.025	0.004	0	0	0	0	1
	men	1384	0.080	0.007	0	0	0	0	1
temporary employed	general	2697	0.263	0.009	0	0	0	1	1
	women	1302	0.442	0.014	0	0	0	1	1
	men	1384	0.095	0.008	0	0	0	0	1
dispatched worker	general	2697	0.021	0.003	0	0	0	0	1
	women	1302	0.031	0.005	0	0	0	0	1
	men	1384	0.012	0.003	0	0	0	0	1
self-employed	general	2697	0.087	0.005	0	0	0	0	1
	women	1302	0.058	0.007	0	0	0	0	1
	men	1384	0.113	0.009	0	0	0	0	1
family worker	general	2697	0.043	0.004	0	0	0	0	1

	women	1302	0.069	0.007	0	0	0	0	1
	men	1384	0.017	0.004	0	0	0	0	1
employed at government agency	general	2697	0.090	0.006	0	0	0	0	1
	women	1302	0.093	0.008	0	0	0	0	1
	men	1384	0.087	0.008	0	0	0	0	1
commuting time	general	2697	2.428	0.050	0	1	2	3	9
(for definitions of categories	women	1302	2.021	0.057	0	1	1	3	9
see Section A of appendix)	men	1384	2.813	0.069	0	1	2	4	9
satisfaction with health	general	2697	2.503	0.019	0	2	2	3	4
	women	1302	2.56	0.028	0	2	3	3	4
	men	1384	2.455	0.027	0	2	2	3	4
satisfaction with area of living	general	2697	2.754	0.020	0	2	3	4	4
	women	1302	2.719	0.029	0	2	3	4	4
	men	1384	2.783	0.027	0	2	3	4	4
satisfaction with leisure	general	2697	2.406	0.019	0	2	2	3	4
	women	1302	2.4	0.029	0	2	2	3	4
	men	1384	2.408	0.027	0	2	2	3	4
member in civil/social organisation	general	2697	0.078	0.005	0	0	0	0	1
	women	1302	0.064	0.007	0	0	0	0	1
	men	1384	0.092	0.008	0	0	0	0	1
member in hobby organization	general	2697	0.113	0.006	0	0	0	0	1
	women	1302	0.125	0.09	0	0	0	0	1
	men	1384	0.101	0.008	0	0	0	0	1

frequency of meeting friends	general	2697	2.538	0.023	0	2	3	3	5
(for definitions of categories	women	1302	2.648	0.032	0	2	3	3	5
see Section A of appendix)	men	1384	2.44	0.033	0	2	2	3	5
frequency of doing sports	general	2697	1.433	0.031	0	0	1	3	4
(for definitions of categories	women	1302	1.274	0.044	0	0	0	3	4
see Section A of appendix)	men	1384	1.58	0.043	0	0	1	3	4
number of books/comics/magazines	general	2697	0.885	0.023	0	0	0	1	4
	women	1302	0.875	0.032	0	0	0	1	4
	men	1384	0.892	0.032	0	0	0	1	4
hours watching TV	general	2697	2.923	0.030	0	2	3	4	$\geq$7
	women	1302	3.075	0.046	0	2	3	4	$\geq$7
	men	1384	2.78	0.039	0	2	3	4	$\geq$7

Table 5: Means and SEs of Variables (Sub-sample A)

		n	mean	SE of mean	min.	1nd qu.	median	3rd qu.	max.
job satisfaction	general	1110	1.828	0.028	0	1	2	3	3
	women	549	1.831	0.039	0	1	2	3	3
	men	561	1.825	0.040	0	1	2	3	3
age	general	1110	45.53	0.369	20	36	46	56	69
	women	549	44.66	0.520	20	35	45	54	69
	men	561	46.37	0.521	20	37	47	57	69
male	general	1110	0.51	0.015	0	0	1	1	1
	women	549	0	0	0	0	0	0	0
	men	561	1	0	1	1	1	1	1
never married	general	1110	0.186	0.012	0	0	0	0	1
	women	549	0.20	0.017	0	0	0	0	1
	men	561	0.171	0.016	0	0	0	0	1
divorced	general	1110	0.06	0.007	0	0	0	0	1
	women	549	0.086	0.012	0	0	0	0	1
	men	561	0.036	0.008	0	0	0	0	1
widowed	general	1110	0.028	0.004	0	0	0	0	1
	women	549	0.046	0.009	0	0	0	0	1
	men	561	0.011	0.004	0	0	0	0	1
number of children	general	1110	1.572	0.033	0	0	2	2	≥4
	women	549	1.563	0.048	0	0	2	2	≥4

	men	561	1.581	0.047	0	0	2	2	≥4
graduated from college	general	1110	0.431	0.015	0	0	0	1	1
	women	549	0.421	0.021	0	0	0	1	1
	men	561	0.440	0.021	0	0	0	1	1
income	general	1110	6.752	0.093	≤2	4	7	9	≥14
(for definitions of categories	women	549	4.996	0.104	≤2	3	5	7	≥13
see Section A of appendix)	men	561	8.471	0.114	≤2	7	8	10	≥14
work hours	general	1110	39.48	0.483	2	30	40	48	112
	women	549	32.62	0.588	2	22	35	40	92
	men	561	46.19	0.647	5	40	45	53	112
executive position	general	1110	0.049	0.006	0	0	0	0	1
	women	549	0.015	0.005	0	0	0	0	1
	men	561	0.082	0.012	0	0	0	0	1
temporary employed	general	1110	0.279	0.013	0	0	0	1	1
	women	549	0.474	0.021	0	0	0	1	1
	men	561	0.095	0.009	0	0	0	0	1
dispatched worker	general	1110	0.024	0.005	0	0	0	0	1
	women	549	0.038	0.008	0	0	0	0	1
	men	561	0.011	0.004	0	0	0	0	1
self-employed	general	1110	0.068	0.008	0	0	0	0	1
	women	549	0.042	0.009	0	0	0	0	1
	men	561	0.093	0.012	0	0	0	0	1
family worker	general	1110	0.031	0.005	0	0	0	0	1

	women	549	0.051	0.009	0	0	0	0	1
	men	561	0.011	0.004	0	0	0	0	1
employed at government agency	general	1110	0.095	0.009	0	0	0	0	1
	women	549	0.102	0.013	0	0	0	0	1
	men	561	0.087	0.012	0	0	0	0	1
commuting time	general	1110	2.415	0.070	0	1	2	3	9
(for definitions of categories	women	549	2.027	0.087	0	1	1	3	9
see Section A of appendix)	men	561	2.795	0.107	0	1	2	4	9
satisfaction with health	general	1110	2.539	0.030	0	2	2	3	4
	women	549	2.587	0.044	0	2	3	3	4
	men	561	2.492	0.042	0	2	2	3	4
satisfaction with area of living	general	1110	2.754	0.031	0	2	3	4	4
	women	549	2.707	0.045	0	2	3	4	4
	men	561	2.8	0.043	0	2	3	4	4
satisfaction with leisure	general	1110	2.45	0.031	0	2	2	3	4
	women	549	2.434	0.044	0	2	2	3	4
	men	561	2.467	0.043	0	2	2	3	4
member in civil/social organisation	general	1110	0.075	0.008	0	0	0	0	1
	women	549	0.062	.0.010	0	0	0	0	1
	men	561	0.088	0.008	0	0	0	0	1
member in hobby organization	general	1110	0.109	0.009	0	0	0	0	1
	women	549	0.113	0.014	0	0	0	0	1
	men	561	0.105	0.012	0	0	0	0	1

frequency of meeting friends	general	1110	2.495	0.035	0	2	2	3	5
(for definitions of categories	women	549	2.612	0.048	0	2	3	3	5
see Section A of appendix)	men	561	2.38	0.050	0	2	2	3	5
frequency of doing sports	general	1110	1.421	0.049	0	0	1	3	4
(for definitions of categories	women	549	1.251	0.069	0	0	0	3	4
see Section A of appendix)	men	561	1.586	0.068	0	0	1	3	4
number of books/comics/magazines	general	1110	0.890	0.035	0	0	0	1	4
	women	549	0.962	0.051	0	0	1	1	4
	men	561	0.82	0.047	0	0	0	1	4
hours watching TV	general	1110	2.897	0.046	0	2	3	4	$\geq$7
	women	549	3.005	0.069	0	2	3	4	$\geq$7
	men	561	2.791	0.061	0	2	3	4	$\geq$7
frequency of private trips	general	1110	1.086	0.024	0	0	1	2	3
	women	549	1.124	0.034	0	1	1	2	3
	men	561	1.05	0.034	0	0	1	2	3
self-reported health	general	1110	2.799	0.031	0	2	3	4	4
	women	549	2.84	0.043	0	2	3	4	4
	men	561	2.759	0.044	0	2	3	4	4
self-reported happiness (A)	general	1110	1.906	0.026	0	1	2	3	3
	women	549	1.893	0.037	0	1	2	3	3
	men	561	1.92	0.037	0	1	2	3	3
progression of economic situation	general	1110	0.574	0.020	0	0	0	1	2
in the last years	women	549	0.623	0.028	0	0	1	1	2
	men	561	0.526	0.027	0	0	0	1	2

believe in opportunities to improve one's life	general	1110	1.628	0.028	0	1	2	2	4
	women	549	1.576	0.039	0	1	2	2	4
	men	561	1.679	0.040	0	1	2	2	4
self-reported social status	general	1110	1.567	0.025	0	1	2	2	4
	women	549	1.536	0.034	0	1	2	2	4
	men	561	1.597	0.036	0	1	2	2	4
anxiety regarding economic future	general	1110	2.941	0.028	0	3	3	4	4
	women	549	2.962	0.041	0	3	3	4	4
	men	561	2.92	0.039	0	3	3	4	4
trust in people and organizations	general	1110	1.125	0.017	0	1	1	1	2
	women	549	1.14	0.022	0	1	1	1	2
	men	561	1.111	0.025	0	1	1	1	2

Table 6: MEANS AND SES OF VARIABLES (SUB-SAMPLE B)

		n	mean	SE of mean	min.	1nd qu.	median	3rd qu.	max.
job satisfaction	general	990	1.814	0.029	0	1	2	3	3
	women	458	1.913	0.043	0	1	2	3	3
	men	532	1.729	0.040	0	1	2	2	3
age	general	990	44.99	0.407	20	35	45	55.75	69
	women	458	43.66	0.586	20	34	44	54	69
	men	532	46.13	0.561	20	36	46	57.25	69
male	general	990	0.537	0.016	0	0	1	1	1
	women	458	0	0	0	0	0	0	0
	men	532	1	0	1	1	1	1	1
never married	general	990	0.218	0.013	0	0	0	0	1
	women	458	0.218	0.019	0	0	0	0	1
	men	532	0.218	0.018	0	0	0	0	1
divorced	general	990	0.046	0.007	0	0	0	0	1
	women	458	0.032	0.011	0	0	0	0	1
	men	532	0.032	0.008	0	0	0	0	1
widowed	general	990	0.013	0.004	0	0	0	0	1
	women	458	0.017	0.006	0	0	0	0	1
	men	532	0.009	0.004	0	0	0	0	1
number of children	general	990	1.492	0.036	0	0	2	2	$\geq$4
	women	458	1.489	0.0529	0	0	2	2	$\geq$4

	men	532	1.494	0.049	0	0	2	2	≥4
graduated from college	general	990	0.429	0.016	0	0	0	1	1
	women	458	0 .419	0.024	0	0	0	1	1
	men	532	0.438	0.022	0	0	0	1	1
income	general	990	7.011	0.102	≤2	4	7	9	≥14
(for definitions of categories	women	458	5.201	0.120	≤2	3	5	7	≥14
see Section A of appendix)	men	532	8.57	0.124	≤2	7	8	11	≥14
work hours	general	990	40.04	0.499	2	30	40	50	105
	women	458	33.22	0.667	2	21	36	42	80
	men	532	45.9	0.627	5	40	45	50	105
executive position	general	990	0.058	0.007	0	0	0	0	1
	women	458	0.033	0.008	0	0	0	0	1
	men	532	0.079	0.012	0	0	0	0	1
temporary employed	general	990	0.262	0.014	0	0	0	1	1
	women	458	0.445	0.023	0	0	0	1	1
	men	532	0.103	0.013	0	0	0	0	1
dispatched worker	general	990	0.016	0.004	0	0	0	0	1
	women	458	0.024	0.007	0	0	0	0	1
	men	532	0.009	0.004	0	0	0	0	1
self-employed	general	990	0.066	0.008	0	0	0	0	1
	women	458	0.041	0.009	0	0	0	0	1
	men	532	0.086	0.012	0	0	0	0	1
family worker	general	990	0.031	0.006	0	0	0	0	1

	women	458	0.046	0.010	0	0	0	0	1
	men	532	0.019	0.006	0	0	0	0	1
employed at government agency	general	990	0.091	0.009	0	0	0	0	1
	women	458	0.092	014	0	0	0	0	1
	men	532	0.090	0.012	0	0	0	0	1
commuting time	general	990	2.54	0.075	0	1	2	4	9
(for definitions of categories	women	458	2.085	0.093	0	1	1	3	9
see Section A of appendix)	men	532	2.932	0.111	0	1	2	4	9
satisfaction with health	general	990	2.466	0.032	0	2	2	3	4
	women	458	2.537	0.046	0	2	2	3	4
	men	532	2.404	0.044	0	2	2	3	4
satisfaction with area of living	general	990	2.758	0.032	0	2	3	4	4
	women	458	2.722	0.032	0	2	3	4	4
	men	532	2.786	0.044	0	2	3	4	4
satisfaction with leisure	general	990	2.358	0.032	0	2	2	3	4
	women	458	2.365	0.048	0	2	2	3	4
	men	532	2.352	0.043	0	2	2	3	4
member in civil/social organisation	general	990	0.075	0.008	0	0	0	0	1
	women	458	0.050	0.048	0	0	0	0	1
	men	532	0.096	0.013	0	0	0	0	1
member in hobby organization	general	990	0.104	0.010	0	0	0	0	1
	women	458	0.116	0.015	0	0	0	0	1
	men	532	0.094	0.013	0	0	0	0	1

frequency of meeting friends	general	990	2.619	0.038	0	2	3	3	5
(for definitions of categories	women	458	2.753	0.054	0	2	3	3	5
see Section A of appendix)	men	532	2.504	0.054	0	2	2	3	5
frequency of doing sports	general	990	1.452	0.051	0	0	1	3	4
(for definitions of categories	women	458	1.271	0.075	0	0	0	3	4
see Section A of appendix)	men	532	1.607	0.070	0	0	1	3	4
number of books/comics/magazines	general	990	0.893	0.038	0	0	0	1	4
	women	458	0.856	0.054	0	0	0	1.25	4
	men	532	0.925	0.054	0	0	0	1	4
hours watching TV	general	990	2.917	0.050	0	2	3	4	≥ 7
	women	458	3.098	0.078	0	2	3	4	≥ 7
	men	532	2.761	0.065	0	2	3	4	≥ 7
lonelyness	general	990	0.963	0.024	0	0	1	1	3
	women	458	0.974	0.035	0	0	1	1	3
	men	532	0.953	0.034	0	0	1	1	3
pain intefering with daily activities	general	990	1.597	0.028	1	1	1	2	5
	women	458	1.686	0.043	1	1	1	2	5
	men	532	1.521	0.036	1	1	1	2	5
self-reported happiness (B)	general	990	3.687	0.028	1	3	4	4	5
	women	458	3.705	0.040	1	3	4	4	5
	men	532	3.671	0.038	0	3	4	4	5
solidarity at work place	general	990	1.546	0.029	0	1	2	2	3
	women	458	1.6	0.042	0	1	2	2	3
	men	532	1.5	0.041	0	1	2	2	3

progression of economic situation	general	990	0.631	0.020	0	0	1	1	2
in the last years	women	458	0.681	0.031	0	0	1	1	2
	men	532	0.588	0.027	0	0	1	1	2
believe in opportunities	general	990	1.495	0.029	0	1	1	2	4
to improve one's life	women	458	1.487	0.042	0	1	1	2	4
	men	532	1.502	0.041	0	1	1	2	4
self-reported social status	general	990	1.646	0.025	0	1	2	2	4
	women	458	1.657	0.036	0	1	2	2	4
	men	532	1.637	0.034	0	1	2	2	4
anxiety regarding economic future	general	990	2.904	0.030	0	2	3	4	4
	women	458	2.928	0.045	0	2	3	4	4
	men	532	2.883	0.040	0	2	3	4	4

C Regression Tables

Table 7: Regression Analysis Main Sample (OLS) - models using income proxy variable

	general	women	men
age	-0.010	-0.023	0.001
	(0.011)	(0.015)	(0.017)
	[65.041]	[66.419]	[75.603]
age^2	0.0001	0.0002	-0.00001
	(0.0001)	(0.0002)	(0.0002)
	[62.791]	[63.203]	[75.139]
male	-0.094^{**}		
	(0.041)		
	[1.417]		
never married	-0.230^{***}	-0.310^{***}	-0.171^{**}
	(0.063)	(0.096)	(0.086)
	[2.191]	[2.422]	[2.141]
divorced	-0.014	-0.092	0.085
	(0.087)	(0.108)	(0.147)
	[1.065]	[1.105]	[1.039]
widowed	0.044	-0.114	0.559^{***}
	(0.114)	(0.131)	(0.193)
	[1.068]	[1.113]	[1.026]
number of children	0.018	0.012	0.027
	(0.022)	(0.032)	(0.030)
	[2.094]	[2.242]	[2.024]
graduated from college	0.092^{**}	0.066	0.110^{**}
	(0.036)	(0.052)	(0.051)
	[1.178]	[1.222]	[1.186]
satisfaction with financial situation	0.138^{***}	0.116^{***}	0.158^{***}
	(0.018)	(0.025)	(0.026)
	[1.336]	[1.346]	[1.369]

work hours (log)	−0.031	−0.064	0.001
	(0.039)	(0.049)	(0.068)
	[1.551]	[1.434]	[1.398]
executive position	0.164**	−0.022	0.191**
	(0.083)	(0.177)	(0.095)
	[1.152]	[1.120]	[1.178]
temporary employed	−0.051	−0.093	−0.065
	(0.052)	(0.065)	(0.098)
	[1.773]	[1.796]	[1.418]
dispatched worker	−0.238*	−0.211	−0.344
	(0.124)	(0.139)	(0.255)
	[1.061]	[1.094]	[1.046]
self-employed	0.050	−0.010	0.072
	(0.068)	(0.122)	(0.081)
	[1.274]	[1.318]	[1.262]
family worker	0.054	−0.038	0.233
	(0.091)	(0.115)	(0.149)
	[1.178]	[1.332]	[1.054]
employed at government agency	0.151***	0.132	0.170**
	(0.057)	(0.081)	(0.082)
	[1.082]	[1.103]	[1.082]
commuting time	0.008	0.018	0.004
	(0.008)	(0.013)	(0.010)
	[1.262]	[1.362]	[1.200]
satisfaction with health	0.051**	0.068**	0.033
	(0.020)	(0.029)	(0.028)
	[1.302]	[1.286]	[1.335]
satisfaction with area of living	0.044**	0.071**	0.019
	(0.020)	(0.028)	(0.029)
	[1.277]	[1.317]	[1.273]
satisfaction with leisure	0.070***	0.053*	0.089***
	(0.022)	(0.032)	(0.029)
	[1.486]	[1.544]	[1.462]

member in civil/social organisation	0.157***	0.109	0.189**
	(0.059)	(0.090)	(0.080)
	[1.075]	[1.073]	[1.090]
member in hobby organization	0.081	0.072	0.113
	(0.052)	(0.069)	(0.078)
	[1.082]	[1.122]	[1.067]
frequency of meeting friends	0.016	0.003	0.025
	(0.016)	(0.025)	(0.021)
	[1.190]	[1.191]	[1.212]
frequency of doing sports	0.007	0.003	0.012
	(0.011)	(0.016)	(0.015)
	[1.113]	[1.102]	[1.128]
number of books/comics/magazines	0.049***	0.037	0.060***
	(0.016)	(0.023)	(0.021)
	[1.155]	[1.158]	[1.186]
hours watching TV	−0.012	−0.008	−0.016
	(0.011)	(0.016)	(0.017)
	[1.156]	[1.149]	[1.180]
Constant	2.336**	4.130***	−0.117
	(0.946)	(1.335)	(1.456)
Observations	2697	1302	1384
R^2	0.12	0.107	0.145
Adjusted R^2	0.111	0.090	0.126
Residual Std. Error	0.874	0.876	0.880
F Statistic	13.95*** (df = 26; 2659)	6.12*** (df = 25; 1038)	7.567*** (df = 25; 1276)
Note:		*p<0.1; **p<0.05; ***p<0.01	

Table 8: Regression Analysis Main Sample (Ordered Logit) - models using income proxy variable

	general	women	men
age	−0.029	−0.057*	−0.012
	(0.023)	(0.033)	(0.035)
	[62.104]	[61.777]	[77.351]
age^2	0.0003	0.001*	0.0001
	(0.0002)	(0.0004)	(0.0004)
	[60.539]	[59.386]	[78.245]
male	−0.189**		
	(0.086)		
	[1.489]		
never married	−0.486***	−0.729***	−0.316*
	(0.132)	(0.207)	(0.179)
	[2.305]	[2.754]	[2.269]
divorced	0.008	−0.187	0.270
	(0.188)	(0.237)	(0.322)
	[1.077]	[1.140]	[1.110]
widowed	0.042	−0.281	1.085**
	(0.252)	(0.289)	(0.521)
	[1.106]	[1.227]	[1.048]
number of children	0.035	−0.004	0.072
	(0.046)	(0.070)	(0.064)
	[2.203]	[2.371]	[2.245]
graduated from college	0.184**	0.117	0.244**
	(0.077)	(0.109)	(0.110)
	[1.172]	[1.186]	[1.219]
satisfaction with financial situation	0.293***	0.251***	0.337***
	(0.039)	(0.054)	(0.059)
	[1.302]	[1.257]	[1.406]
work hours (log)	−0.081	−0.147	−0.008
	(0.087)	(0.109)	(0.152)
	[1.665]	[1.490]	[1.524]

executive position	0.448**	−0.050	0.540**
	(0.199)	(0.433)	(0.225)
	[1.109]	[1.143]	[1.109]
temporary employed	−0.122	−0.201	−0.177
	(0.110)	(0.139)	(0.207)
	[1.868]	[1.872]	[1.464]
dispatched worker	−0.547**	−0.454	−0.864
	(0.256)	(0.281)	(0.607)
	[1.109]	[1.188]	[1.051]
self-employed	0.088	−0.055	0.141
	(0.147)	(0.274)	(0.176)
	[1.281]	[1.475]	[1.238]
family worker	0.084	−0.114	0.391
	(0.199)	(0.255)	(0.306)
	[1.166]	[1.321]	[1.115]
employed at government agency	0.327**	0.280	0.393**
	(0.128)	(0.180)	(0.184)
	[1.097]	[1.126]	[1.111]
commuting time	0.014	0.032	0.008
	(0.017)	(0.029)	(0.021)
	[1.283]	[1.531]	[1.211]
satisfaction with health	0.115***	0.149**	0.083
	(0.044)	(0.063)	(0.062)
	[1.355]	[1.316]	[1.402]
satisfaction with area of living	0.102**	0.176***	0.040
	(0.043)	(0.060)	(0.062)
	[1.298]	[1.344]	[1.339]
satisfaction with leisure	0.159***	0.114	0.209***
	(0.047)	(0.070)	(0.064)
	[1.470]	[1.620]	[1.417]
member in civil/social organization	0.306**	0.205	0.386**
	(0.132)	(0.199)	(0.179)
	[1.120]	[1.120]	[1.156]

member in hobby organization	0.171	0.129	0.269
	(0.114)	(0.150)	(0.175)
	[1.142]	[1.225]	[1.114]
frequency of meeting friends	0.036	0.006	0.057
	(0.034)	(0.054)	(0.045)
	[1.230]	[1.265]	[1.254]
frequency of doing sports	0.015	0.012	0.021
	(0.024)	(0.034)	(0.033)
	[1.146]	[1.100]	]1.210]
number of books/comics/magazines	0.104***	0.083*	0.122***
	(0.034)	(0.049)	(0.046)
	[1.189]	[1.229]	[1.192]
hours of watching TV	−0.026	−0.020	−0.034
	(0.024)	(0.034)	(0.036)
	[1.172]	[1.204]	[1.188]
Constants:			
y>=1	4.090**	8.099***	−0.449
	(2.013)	(2.901)	(3.236)
y>=2	2.447	6.463**	−2.120
	(2.008)	(2.892)	(3.229)
y>=3	0.524	4.534	−4.063
	(2.007)	(2.888)	(3.227)
Observations	2697	1302	1385
McFadden R^2	0.052	0.046	0.063
χ^2	355.85***	152.80***	224.39***
	(df = 26)	(df = 25)	(df = 25)
Log-likelihood	-3277.514	-1583.715	-1681.388
	(df=29)	(df=28)	(df=28)

Table 9: REGRESSION ANALYSIS MAIN SAMPLE (OLS) - MODELS WITHOUT IRRELEVANT VARIABLE

	general	women	men
age	-0.031^{***}	-0.030^{*}	-0.035^{*}
	(0.012)	(0.017)	(0.019)
	[64.003]	[62.181]	[76.448]
age^2	0.0003^{***}	0.0003^{*}	0.0004^{**}
	(0.0001)	(0.0002)	(0.0002)
	[61.460]	[58.502]	[75.820]
male	-0.258^{***}		
	(0.047)		
	[1.539]		
never married	-0.261^{***}	-0.347^{***}	-0.206^{***}
	(0.056)	(0.087)	(0.077)
	[1.499]	[1.846]	[1.437]
divorced		-0.134	
		(0.116)	
		[1.073]	
widowed			0.563^{***}
			(0.178)
			[1.026]
graduated from college	0.078^{**}	0.082	0.063
	(0.039)	(0.056)	(0.054)
	[1.140]	[1.164]	[1.147]
income	0.059^{***}	0.042^{***}	0.070^{***}
	(0.009)	(0.016)	(0.012)
	[1.986]	[2.004]	[1.540]
work hours (log)	-0.122^{***}	-0.133^{**}	-0.088
	(0.043)	(0.058)	(0.069)
	[1.509]	[1.542]	[1.269]
temporary employed		-0.081	
		(0.074)	
		[1.873]	

dispatched worker	-0.272^{**}	-0.231	-0.436^{*}
	(0.135)	(0.162)	(0.257)
	[1.023]	[1.107]	[1.023]
family worker		-0.134	0.424^{**}
		(0.139)	(0.171)
		[1.184]	[1.044]
employed at government agency	0.106^{*}		0.144
	(0.063)		(0.089)
	[1.061]		[1.055]
commuting time		0.015	
		(0.014)	
		[1.317]	
satisfaction with health	0.070^{***}	0.076^{**}	0.063^{**}
	(0.022)	(0.031)	(0.030)
	[1.244]	[1.222]	[1.263]
satisfaction with area of living	0.059^{***}	0.092^{***}	0.032
	(0.021)	(0.030)	(0.030)
	[1.240]	[1.260]	[1.245]
satisfaction with leisure	0.099^{***}	0.071^{**}	0.133^{***}
	(0.023)	(0.033)	(0.031)
	[1.370]	[1.375]	[1.369]
member in civil/social organization	0.159^{**}	0.165^{*}	0.156^{*}
	(0.066)	(0.099)	(0.088)
	[1.047]	[1.059]	[1.044]
member in hobby organization	0.087	0.090	0.110
	(0.057)	(0.078)	(0.085)
	[1.068]	[1.096]	[1.056]
frequency doing sports	0.012		
	(0.012)		
	[1.092]		
number of books/comics/magazines	0.056^{***}	0.038	0.074^{***}
	(0.017)	(0.024)	(0.023)
	[1.120]	[1.118]	[1.139]

Constant	3.575***	4.827***	0.238
	(0.547)	(1.121)	(1.324)
Observations	2209	1064	1145
R^2	0.112	0.089	0.142
Adjusted R^2	0.105	0.075	0.130
Residual Std. Error	0.877	0.875	0.878
F Statistic	17.21*** (df = 16; 2192)	6.389*** (df = 16; 1047)	11.69*** (df = 16; 1128)
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 10: Regression Analysis Subsample A Sample (OLS) - models without irrelevant variable

	general	women	men
male	−0.093		
	(0.065)		
	[1.621]		
never married	−0.199***	−0.364***	
	(0.069)	(0.126)	
	[1.219]	[2.008]	
divorced	0.125		
	[1.084]		
	(0.114)		
widowed	0.175		0.594*
	(0.152)		(0.342)
	[1.054]		[1.044]
number of children		−0.050	
		(0.046)	
		[2.102]	
income	0.037***	0.023	0.040***
	(0.012)	(0.021)	(0.014)
	[2.074]	[2.056]	[1.236]
work hours (log)	−0.160***	−0.200**	
	(0.059)	(0.088)	
	[1.517]	[1.602]	
executive position	0.131	−0.344	0.190
	(0.117)	(0.288)	(0.124)
	[1.085]	[1.052]	[1.080]
temporary employed		−0.106	
		(0.101)	
		[1.990]	
family worker	−0.249	−0.440**	
	(0.158)	(0.182)	
	[1.036]	[1.175]	

employed at government agency		0.123	
		(0.114)	
		[1.071]	
satisfaction with area of living	0.050*	0.091**	
	(0.029)	(0.037)	
	[1.332]	[1.172]	
satisfaction with leisure	0.067**		0.110***
	(0.031)		(0.042)
	[1.511]		[1.506]
member in civil/social oraganization	0.152*	0.244*	
	(0.092)	(0.142)	
	[1.065]	[1.077]	
frequency of meeting friends	−0.030		−0.037
	(0.025)		(0.033)
	[1.217]		]1.114]
number of books/comics/magazines	0.036		0.066**
	(0.022)		(0.030)
	[1.078]		[1.127]
frequency of private trips	0.070**		0.101**
	(0.034)		(0.048)
	[1.263]		[1.252]
self-reported health	0.030		0.066*
	(0.027)		(0.034)
	[1.275]		[1.331]
self-reported happiness (A)	0.087**	0.149***	0.092*
	(0.037)	(0.047)	(0.052)
	[1.740]	[1.315]	[1.834]
progression of economic situation in last years	0.132***	0.085	
	(0.043)	(0.062)	
	[1.209]	[1.237]	
believe in opportunities to improve one's life	0.055*	0.117***	
	(0.029)	(0.041)	
	[1.164]	[1.150]	

self-reported social status	0.120***		0.186***
	(0.036)		(0.049)
	[1.498]		[1.529]
anxiety regarding economic future	−0.106***	−0.171***	−0.094**
	(0.031)	(0.040)	(0.040)
	[1.414]	[1.314]	]1.239]
trust in people and organizations	0.061	−0.091	0.162**
	(0.049)	(0.075)	(0.063)
	[1.101]	[1.096]	[1.110]
Constant	2.882***	5.938***	−0.629
	(0.969)	(1.230)	(0.727)
Observations	1110	549	561
R^2	0.208	0.185	0.256
Adjusted R^2	0.193	0.162	0.240
Residual Std. Error	0.832	0.831	0.824
F Statistic	13.6*** (df = 21; 1088)	8.078*** (df = 15; 533)	15.72*** (df = 12; 548)
Note:	*p<0.1; **p<0.05; ***p<0.01		

Table 11: Regression Analysis Subsample B Sample (OLS) - models without irrelevant variable

	general	women	men
age^2	0.0001^{***}	0.0002^{***}	0.0001^{***}
	(0.00003)	(0.00004)	(0.00004)
	[1.450]	[1.158]	[1.455]
male	-0.281^{***}		
	(0.065)		
	[1.495]		
never married	−0.107		−0.117
	(0.077)		(0.108)
	[1.463]		[1.594]
widowed			0.213^{*}
			(0.127)
			[1.021]
graduated from college	0.099^{*}	0.205^{**}	
	(0.056)	(0.080)	
	[1.145]	[1.102]	
income	0.035^{***}		0.032^{**}
	(0.012)		(0.016)
	[2.139]		[1.470]
work hours (log)	−0.092		
	(0.062)		
	[1.580]		
executive position	−0.178		-0.272^{*}
	(0.140)		(0.161)
	[1.123]		[1.206]
temporary		−0.089	
		(0.081)	
		[1.118]	
dispatched	−0.340		-0.676^{**}
	(0.231)		(0.307)
	[1.031]		[1.041]

self-employed		0.348*	−0.190
		(0.209)	(0.120)
		[1.099]	
family worker	0.210		0.448*
	(0.129)		(0.236)
	[1.057]		[1.064]
commuting time	0.013	0.042**	
	(0.012)	(0.019)	
	[1.230]	[1.111]	
satisfaction with health		0.118***	
		(0.044)	
		[1.246]	
satisfaction with leisure	0.038		0.046
	(0.032)		(0.043)
	[1.237]		[1.250]
member in civil/social organization			0.182
			(0.124)
			[1.101]
frequency of meeting friends		−0.041	
		(0.037)	
		[1.110]	
number of books/comics/magazines	0.049**		0.074**
	(0.023)		(0.031)
	[1.126]		[1.063]
loneliness	−0.054	−0.148***	
	(0.041)	(0.058)	
	[1.239]	[1.203]	
pain interfering with daily activities	−0.069**		−0.091**
	(0.032)		(0.042)
	[1.058]		[1.094]
self-reported happiness (B)	0.080**	0.070	0.074
	(0.038)	(0.053)	(0.051)
	[1.487]	[1.414]	[1.527]

solidarity at work place	0.171***	0.203***	0.160***
	(0.031)	(0.046)	(0.040)
	[1.094]	[1.077]	[1.087[
progression of economic situation in last years	0.163***	0.119*	0.194***
	(0.047)	(0.066)	(0.063)
	[1.275]	[1.172]	[1.305]
believe in opportunities to improve one's life	0.085***	0.072	0.088**
	(0.032)	(0.046)	(0.044)
	[1.155]	[1.116]	[1.230]
self-reported social status			0.086
			(0.056)
			[1.515]
anxiety regarding economic future	−0.055*		−0.062
	(0.031)		(0.044)
	[1.336]		[1.395]
Constant	1.447	−0.587	0.924
	(1.033)	(0.966)	(1.783)
Observations	990	458	532
R^2	0.191	0.188	0.226
Adjusted R^2	0.176	0.166	0.199
Residual Std. Error	0.841	0.831	0.834
F Statistic	12.08*** (df = 19; 970)	8.582*** (df = 12; 445)	8.304*** (df = 18, 513)
Note:			*p<0.1; **p<0.05; ***p<0.01

D Correlation Matrix

Table 12: Correlations (Main Sample)

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
1. Age	g.	1.00																									
	w.	1.00																									
	m.	1.00																									
2. Age²	g.	0.99	1.00																								
	w.	0. 99	1.00																								
	m.	0.99	1.00																								
3. Male	g.	0.08	0.08	1.00																							
4. Never married	g.	-0.49	-0.44	-0.01	1.00																						
	w.	-0.54	-0.49	NA	1.00																						
	m.	-0.44	-0.41	NA	1.00																						
5. Divorced	g.	0.06	0.05	-0.10	-0.12	1.00																					
	w.	0.07	0.06	NA	-0.15	1.00																					
	m.	0.07	0.06	NA	-0.09	1.00																					
6. Widowed	g.	0.17	0.18	-0.08	-0.07	-0.04	1.00																				
	w.	0.23	0.25	NA	-0.09	-0.05	1.00																				
	m.	0.12	0.13	NA	-0.05	-0.02	1.00																				
7. No. of children	g.	0.51	0.48	-0.001	0.68	0.02	0.10	1.00																			
	w.	0.56	0.53	NA	-0.69	0.04	0.13	1.00																			
	m.	0.46	0.44	NA	-0.67	0.004	0.07	1.00																			
8. College graduation	g.	-0.15	-0.16	0.02	0.07	-0.07	-0.05	-0.12	1.00																		
	w.	-0.21	-0.23	NA	0.14	-0.08	-0.06	-0.18	1.00																		
	m.	-0.09	-0.11	NA	0.001	-0.06	-0.04	-0.07	1.00																		
9. Income	g.	0.02	-0.02	0.54	-0.04	-0.02	-0.05	0.001	0.20	1.00																	
	w.	-0.11	-0.11	NA	0.22	0.09	-0.01	-0.23	0.20	1.00																	
	m.	0.03	-0.04	NA	-0.26	-0.02	0.005	0.19	0.24	1.00																	
10. Work hours	g.	-0.14	-0.16	0.36	0.09	0.02	-0.09	-0.12	0.09	0.53	1.00																
	w.	-0.19	-0.19	NA	0.24	0.09	-0.08	-0.24	0.07	0.53	1.00																
	m.	-0.17	-0.22	NA	-0.07	0.02	-0.03	0.01	0.12	0.33	1.00																
11. Executive position	g.	0.14	0.13	0.13	-0.10	-0.02	0.02	0.08	0.05	0.22	0.10	1.00															
	w.	0.11	0.12	NA	-0.06	0.01	0.04	0.04	-0.02	0.06	-0.02	1.00															
	m.	0.14	0.14	NA	-0.12	-0.02	0.03	0.10	0.08	0.24	0.12	1.00															
12. Temp. employed	g.	0.07	0.09	-0.41	-0.08	-0.01	0.05	0.10	-0.11	-0.60	-0.52	-0.14	1.00														
	w.	0.11	0.11	NA	-0.22	-0.08	0.01	0.23	-0.12	-0.57	-0.46	-0.14	1.00														
	m.	0.13	0.18	NA	0.10	0.01	0.05	-0.07	-0.11	-0.45	-0.41	-0.10	1.00														
13. Disp. worker	g.	-0.04	-0.04	-0.07	0.07	0.08	0.02	-0.06	-0.03	-0.06	0.01	-0.03	-0.09	1.00													
	w.	-0.08	-0.08	NA	0.10	0.07	0.03	-0.12	0.06	0.01	0.07	-0.03	-0.16	1.00													

	m.	0.04	0.04	NA	0.04	0.08	-0.01	0.03	-0.02	-0.10	-0.01	-0.03	-0.03	1.00													
14. Self-emp.	g.	0.20	0.21	0.10	-0.08	0.04	0.03	0.10	-0.06	0.01	-0.03	-0.07	-0.18	-0.04	1.00												
	w.	0.19	0.20	NA	-0.07	0.12	0.07	0.08	-0.03	0.001	-0.09	-0.04	-0.22	-0.04	1.00												
	m.	0.21	0.21	NA	-0.09	-0.02	0.02	0.11	-0.09	-0.09	-0.06	-0.10	-0.11	-0.04	1.00												
15. Family	g.	0.06	0.06	-0.11	-0.01	-0.03	-0.01	0.03	-0.03	-0.12	-0.05	-0.04	-0.11	-0.03	-0.06	1.00											
worker	w.	0.15	0.16	NA	-0.10	-0.07	-0.02	0.12	-0.04	-0.05	-0.02	-0.04	-0.22	-0.04	-0.06	1.00											
	m.	-0.08	-0.07	NA	0.14	0.02	-0.01	-0.12	0.01	-0.13	0.03	-0.04	-0.04	-0.01	-0.04	1.00											
16. Emp. at	g.	0.01	-0.004	-0.01	-0.04	-0.01	-0.003	0.01	0.15	0.14	0.04	-0.07	-0.004	-0.02	-0.10	-0.06	1.00										
gov. ag.	w.	0.002	-0.01	NA	-0.02	-0.02	-0.01	-0.02	0.17	0.18	0.08	-0.05	-0.02	-0.02	-0.08	-0.08	1.00										
	m.	0.02	0.01	NA	-0.06	0.01	-0.001	0.04	0.14	0.16	0.004	-0.09	0.01	-0.03	-0.11	-0.04	1.00										
17. Com.	g.	-0.13	-0.13	0.17	0.13	-0.01	-0.08	-0.18	0.17	0.29	0.15	-0.07	-0.16	0.08	-0.19	-0.16	0.08	1.00									
time	w.	-0.25	-0.24	NA	0.33	-0.004	-0.09	-0.35	0.19	0.31	0.21	-0.10	-0.14	0.18	-0.16	-0.19	0.10	1.00									
	m.	-0.08	-0.09	-0.02	NA	0.02	-0.05	-0.05	0.16	0.19	-0.01	-0.10	-0.07	0.02	-0.24	-0.10	0.08	1.00									
18. Sat. with	g.	-0.06	-0.04	-0.06	-0.01	-0.01	-0.01	0.01	0.05	-0.05	-0.07	0.01	0.06	0.01	-0.02	-0.01	0.01	-0.01	1.00								
health	w.	-0.06	-0.05	NA	0.03	-0.03	-0.002	-0.004	0.09	-0.06	-0.06	0.01	0.03	0.03	0.01	-0.02	0.01	0.02	1.00								
	m.	-0.05	-0.03	NA	-0.04	0.002	-0.03	0.03	0.01	0.01	-0.04	0.03	0.04	-0.04	-0.03	0.003	0.02	-0.02	1.00								
19. Sat. with	g.	0.02	0.03	0.03	-0.02	-0.02	0.02	0.04	0.05	0.05	-0.004	0.02	-0.05	0.03	-0.01	-0.004	0.03	0.03	0.28	1.00							
living area	w.	0.002	0.02	NA	0.05	-0.01	0.06	-0.01	0.04	0.04	-0.01	0.03	-0.09	0.06	0.01	0.04	0.02	0.02	0.28	1.00							
	m.	0.03	0.03	NA	-0.09	-0.02	-0.04	0.10	0.06	0.05	-0.02	0.01	0.04	-0.02	-0.04	-0.07	0.03	0.03	0.29	1.00							
20. Sat. with	g.	-0.02	-0.01	0.01	0.03	-0.03	0.01	-0.05	0.01	0.03	-0.05	-0.01	0.002	0.05	0.01	-0.02	0.03	0.01	0.39	0.41	1.00						
leisure	w.	-0.07	-0.06	NA	0.10	-0.04	0.03	-0.12	0.01	0.02	-0.04	0.02	-0.03	0.10	0.03	-0.02	0.01	0.01	0.37	0.41	1.00						
	m.	0.02	0.03	NA	-0.04	-0.02	-0.01	0.02	0.02	0.04	-0.09	-0.04	0.05	-0.01	-0.001	-0.04	0.04	0.003	0.40	0.41	1.00						
21. Member	g.	0.08	0.09	0.05	-0.06	-0.05	0.04	0.08	0.07	0.03	-0.06	0.11	-0.01	-0.04	0.06	0.03	0.04	-0.04	0.04	0.06	0.004	1.00					
of civil or	w.	0.06	0.05	NA	-0.07	-0.05	0.06	0.08	0.08	-0.05	-0.09	0.04	0.02	-0.05	0.04	0.06	0.07	-0.07	0.03	0.03	-0.01	1.00					
social org.	m.	0.10	0.10	NA	-0.06	-0.04	0.03	0.08	0.06	0.03	-0.07	0.14	0.004	-0.03	0.06	0.01	0.03	-0.04	0.05	0.07	0.01	1.00					
22. Mem. of	g.	0.11	0.11	-0.03	-0.04	-0.02	0.05	0.04	0.04	0.002	-0.09	0.03	0.003	-0.01	0.05	0.04	0.06	-0.02	0.05	0.05	0.10	0.13	1.00				
hobby org.	w.	0.16	0.17	NA	-0.05	-0.04	0.07	0.03	0.06	0.03	-0.07	0.02	-0.03	-0.02	0.10	0.06	0.09	-0.04	0.08	0.04	0.11	0.14	1.00				
	m.	0.06	0.07	NA	-0.03	0.004	-0.01	0.05	0.02	0.02	-0.10	0.05	0.03	-0.01	0.02	-0.02	0.03	0.01	0.01	0.05	0.09	0.12	1.00				
23. Meeting	g.	-0.19	-0.17	-0.09	0.21	-0.01	0.02	-0.16	0.11	0.03	-0.07	0.06	0.002	0.002	-0.01	-0.02	-0.01	0.07	0.13	0.03	0.12	0.11	0.12	1.00			
friends	w.	-0.18	-0.16	NA	0.27	-0.03	0.02	-0.20	0.11	0.05	-0.03	0.05	-0.06	0.06	0.03	-0.06	-0.01	0.10	0.14	0.03	0.19	0.10	0.13	1.00			
	m.	-0.20	-0.17	NA	0.16	-0.01	0.01	-0.13	0.12	0.12	-0.06	0.09	-0.01	-0.09	-0.02	0.03	-0.01	0.08	0.11	0.03	0.06	0.13	0.10	1.00			
24. Freq.	g.	0.05	0.05	0.09	0.003	-0.05	0.04	0.02	0.07	0.08	-0.06	0.04	-0.02	0.001	-0.003	0.002	0.07	0.04	0.13	0.06	0.15	0.07	0.11	0.15	1.00		
doing sports	w.	0.08	0.08	NA	0.02	-0.05	0.08	0.03	0.05	-0.02	-0.08	0.02	-0.002	0.03	0.08	0.02	0.02	-0.01	0.13	0.04	0.13	0.06	0.14	0.12	1.00		
	m.	0.01	0.02	NA	-0.01	-0.04	0.01	0.002	0.09	0.09	-0.13	0.04	0.07	-0.02	-0.07	0.01	0.11	0.04	0.15	0.07	0.18	0.06	0.08	0.19	1.00		
25. No. of	g.	0.01	0.01	-0.01	0.09	-0.03	0.01	-0.08	0.19	0.09	-0.02	0.04	-0.05	0.01	0.02	0.003	0.13	0.14	0.06	0.03	0.08	0.10	0.14	0.14	0.17	1.00	
books, comics	w.	0.04	0.03	NA	0.10	-0.02	0.02	-0.09	0.14	0.09	-0.02	0.03	-0.11	0.04	0.08	0.002	0.10	0.14	0.06	0.01	0.11	0.12	0.14	0.09	0.19	1.00	
or magazines	m.	-0.01	-0.01	NA	0.08	-0.048	0.004	-0.08	0.24	0.14	-0.02	0.05	0.02	-0.06	-0.02	0.001	0.15	0.15	0.05	0.04	0.05	0.08	0.15	0.18	0.16	1.00	
26. Hours	g.	0.15	0.16	-0.08	-0.03	0.03	0.07	0.02	-0.17	-0.20	-0.17	0.03	0.15	0.02	0.02	0.03	-0.06	-0.16	-0.04	-0.02	-0.01	-0.05	-0.06	-0.07	-0.04	-0.17	1.00
of TV	w.	0.12	0.14	NA	-0.06	-0.01	0.10	0.03	-0.13	-0.18	-0.13	-0.002	0.14	-0.02	-0.03	0.01	-0.09	-0.18	-0.03	-0.02	-0.05	-0.05	-0.07	-0.07	-0.03	-0.15	1.00
	m.	0.19	0.20	NA	-0.01	0.06	-0.003	0.008	-0.21	-0.18	-0.17	-0.02	0.12	0.06	0.08	0.05	-0.03	-0.13	-0.05	-0.02	0.02	-0.05	-0.05	-0.09	-0.04	-0.18	1.00

E Statistical Tools and Concepts

E.1 Sources of Endogeneity

There are three main sources of endogeneity (Reichstein, n.d.):

1. Omitted Variables: An omitted variable can be defined as *an important explanatory variable that has been left out of a regression equation* (Studenmund, 2005, p. 163). In order to cause an endogeneity bias, the omitted variable must also correlate with one or more of the included independent variables. A famous example for a biased estimation caused by an omitted variable is the significant positive relationship between the number of firefighters at a fire and the amount of property damage, if no control variable measuring the size of the fire is included in the equation ("Testing for Nonspuriousness", n.d.). In the case of our models, possible omitted variables might be certain personality traits,[50] such as having an optimistic attitude towards life or being an extravert.[51]

2. Measurement Error: There are two types of measurement error: Random errors (or white noise errors) that are unrelated to explanatory variables, and systematic errors that are correlated with the explanatory variables (Stutzer & Frey, 2010, p. 9). White noise errors do not cause endogeneity or biased estimates, but reduce the statistical fit by increasing the variability around the average (e.g Stutzer & Frey, 2010, p. 9; Trochim, 2006). It is important to mention this difference, because white noise measurement errors are nearly always an issue in subjective-wellbeing studies.[52] However, systematic measurement errors can become an issue too, especially when several well-being measures are included

[50]In happiness research it is assumed that personality accounts for the majority of the variance between individuals in reported satisfaction (Graaf-Zijl, 2011, p. 204).

[51]Extraversion positively relates to job as well as leisure satisfaction (e.g. Lu & Hu, 2005: Frey, 2008, p. 24; Templer, 2012, p. 117).

[52]E.g. mood variability, most context effects and different measurement reliability of inpendent variables fall in the random error category (E.g. Headey & Wearing, n.d., p. 91; Stutzer & Frey, 2010, p. 9). This type of measruement errors are also prevalent in our models. For example, age is presumably almost perfectly measured, while latent variables such as income and all subjective wellbeing variables are measured with higher error. This might lead to an underestimation of the significance of the realativly poorly measured variables (Headey & Wearing, n.d., p. 91).

in the equation - as is the case for our models. The reason for this is common method variance - meaning the correlation between variables that is unintentionally created through the use of the same method (in our case a questionnaire) to measure each variable (Craighead, Ketchen, Dunn, & Hult, 2011, p. 578). This problem occurs, because people unconsciously tend to adjust their ratings based on previous answers in order to maintain cognitive consistency (Antonakis et al., 2014, p. 22). For example, people who rated different aspects of their life positively (e.g. satisfaction with health, leisure or housing), might feel compelled to evaluate their overall happiness more favorably than they would have done, if no evaluations on different areas of life had been requested. Indeed, there is some evidence that systematic measurement error can arise, if questions on the housing situation precede the question on the general happiness level[53] (Stutzer & Frey, 2010, p. 9).

3. Simultaneity: Simultaneity happens, if two or more variables simultaneously determine each other (e.g. Studenmund, 2005, p. 444; Antonakis et al., 2014, p. 27). It is very common source of endogeneity bias in happiness research, since many important variables are known to represent a determinant as well as the consequence of certain subjective well-being variables (Frey, 2008, p. 11). For instance, in the literature review of the thesis it was mentioned that high job satisfaction contributes to higher happiness levels, but the degree of overall happiness has even more influence on job satisfaction. Similarly, being unemployed surely adds to unhappiness, but unhappy people are also less likely to find employment. Finally, marriage may positively influence life satisfaction, but happier people also have an easier time to find a partner (Stutzer & Frey, 2006). Or as economist Justin Wolfers puts it: *"If you're grumpy who the hell wants to marry you (Freakonomics Blog, 2014)?"*

[53]This problem also arises for our data set, since both questionnaires of the JGSS, ask the respondent to give a rating on his satisfaction with his neighborhood right before question on his general happines level.

E.2 McFadden Pseudo R^2

The terms *predictive efficacy* or *goodness of fit* refer to the amount by which prediction error is reduced thanks to the use of a model (DeMaris, 2012, p. 963). In linear regression, this is assessed with the coefficient of determination (R^2), which is the ratio of the explained sum of squares to the total sum of squares (Studenmund, 2005, p. 50):

$$R^2 = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS} = 1 - \frac{\sum e_i^2}{\sum \left(Y_i - \bar{Y}\right)} \tag{4}$$

Where ESS is the explained sum of squares, both TSS and $\sum \left(Y_i - \bar{Y}\right)$ are the total sum of squares and, both RSS and $\sum e_i^2$ are the residual sum of squares. Unfortunately, there is no commonly accepted analogue in logistic regression, but several competing measures (Hu, Shao, & Palta, 2006). In this thesis the *"likelihood-ratio index"* (also known as $McFadden - Pseudo - R^2$) is used, because it is a very common measure and can be easily calculated in the R software environment. It is based on the classic article of McFadden (1973), where he suggests to compare a model without any predictor to the logit model including all predictor variables. So the *"likelihood-ratio index"* is defined as one minus the ratio of the log likelihood of the model including only the intercept, and the log likelihood of the model with all predictors:

$$McFadden - R^2 = 1 - \left(\frac{lnL(\hat{\theta})}{lnL(\tilde{\theta})}\right) \tag{5}$$

Where $lnL(\hat{\theta})$ is the estimated log likelihood of the full model with all predictors and $lnL(\tilde{\theta})$ is the estimated log likelihood of the intercept model where the slope parameters are all zero.

F Source Code

```
library("Hmisc")
library("MASS")
library("foreign")
library("stargazer")
library("ordinal")
library("VGAM")
library("car")
library("zoo")
library("lmtest")
library("sandwich")
library("AER")
library("systemfit")
library("ivlewbel")
library("mvtnorm")
library("graphics")
library("lmtest")
library("Matrix")
library("robustbase")
library("ivpack")
library("rms")
library("sciplot")

#1. Models using the "respondents income" variable
js_data1 <- happy_data

#Job satisfaction
table(js_data1$st5job)
#Remove missing values and na's
js_data1 <- subset(js_data1, js_data1$st5job != 6 & js_data1$st5job != 9 & js_data1$st5job != 8)

#Age
table(js_data1$ageb)
js_data1$ageb[js_data1$ageb >= 70 ] <- 70
js_data1 <- subset(js_data1, js_data1$ageb != 70)
js_data1$agebSquare <- js_data1$ageb ^ 2

#Income
table(js_data1$szincomx)
#Remove na's, missing values and "no wage"
js_data1 <- subset(js_data1, js_data1$szincomx != 20 & js_data1$szincomx != 21 & js_data1$szincomx != 99 & js_data1$szincomx != 1)
#Combining 15,16,17,18 and 19 to the category "above 10 million yen" due to very few values in the categories
js_data1$szincomx[js_data1$szincomx == 15 | js_data1$szincomx == 16 | js_data1$szincomx == 17 | js_data1$szincomx == 18 | js_data1$szincomx == 19] <- 14
table(js_data1$szincomx)
js_data1$wage <- js_data1$szincomx
```

```
table(js_data1$wage)

#Gender
table(js_data1$sexa)
js_data1$sexa[js_data1$sexa == 2] <- 0
#Coding as dummy vaiable
js_data1$male <- js_data1$sexa
table(js_data1$male)

#Education
table(js_data1$xxlstsch)
#Removing na's, "higher school or vocational school in the old system," "vocational school/commerce school in the old system" due to lack of observation and camparability
js_data1<- subset(js_data1, js_data1$xxlstsch != 99 & js_data1$xxlstsch != 6 & js_data1$xxlstsch != 5 & js_data1$xxlstsch != 4 & js_data1$xxlstsch != 14)
#Combining "ordinary elementary school in the old system", "higher elementary school in the old system", "junior high school/girls' high school in the old system," "junior high school" and "high school" to "high School and below" variable
js_data1$xxlstsch [js_data1$xxlstsch == 1 | js_data1$xxlstsch == 2 | js_data1$xxlstsch == 3 | js_data1$xxlstsch == 8 | js_data1$xxlstsch == 9 ] <- 0
#Combining "university/graduate school in the old system", "university" and "graduate school" to independant variable "college and higher"
js_data1$xxlstsch [js_data1$xxlstsch == 7 | js_data1$xxlstsch == 12 | js_data1$xxlstsch == 13 | js_data1$xxlstsch == 10 | js_data1$xxlstsch == 11 ] <- 1
#Code dummy graduation variable
js_data1$dolstsch [js_data1$dolstsch == 2 | js_data1$dolstsch == 3 | js_data1$dolstsch == 9] <- 0
js_data1$graduatedCollege <- js_data1$dolstsch * js_data1$xxlstsch
table(js_data1$graduatedCollege)

#Marital status
table(js_data1$domarry)
#Removing cohabiting and seperated people from the sample due lack of enough observations
js_data1 <- subset(js_data1, js_data1$domarry != 5 & js_data1$domarry != 6)
#Create dummy variables
js_data1$never_married <- js_data1$domarry
table(js_data1$never_married)
js_data1$never_married [js_data1$never_married == 1 | js_data1$never_married == 2 ] <- 3
js_data1$divorced <- js_data1$domarry
js_data1$divorced [js_data1$divorced == 1 | js_data1$divorced == 3 | js_data1$divorced == 4] <- 1
table(js_data1$divorced)
js_data1$widowed <- js_data1$domarry
```

```
js_data1$widowed [js_data1$widowed == 1 | js_data1$widowed == 4 ] <- 2
table(js_data1$widowed)

#Number of children
table(js_data1$ccnumttl)
#Removing na's
js_data1<- subset(js_data1, js_data1$ccnumttl != 999)
#Combining 4,5,6,7,8 to category "4 and more children" due to small amount of observations
js_data1$ccnumttl [js_data1$ccnumttl == 5 | js_data1$ccnumttl == 6 | js_data1$ccnumttl == 7 | js_data1$ccnumttl == 8] <- 4
table(js_data1$ccnumttl)

#Commuting
#Variable doct: yes, working and living place same, depends
table(js_data1$docmt)
#Eraese na's and "depends" answers
js_data1<- subset(js_data1, js_data1$docmt != 9 & js_data1$docmt != 3)
#Commuting time
table(js_data1$szcmttl)
#Creating new variable by combining both commuting variables. "Identical living and working place" was set to 0.
js_data1$szcmttl[js_data1$docmt == 2] <- 0
#Remove Na's and outliers
js_data1 <- subset(js_data1, js_data1$szcmttl != 9999)
#Combine 0-9 minutes
js_data1$szcmttl[js_data1$szcmttl == 1 | js_data1$szcmttl == 2 | js_data1$szcmttl == 3 | js_data1$szcmttl == 4 | js_data1$szcmttl == 5 | js_data1$szcmttl == 6 | js_data1$szcmttl == 7 | js_data1$szcmttl == 8 | js_data1$szcmttl == 9 ] <- 0
#Combine 10-19 minutes
js_data1$szcmttl[js_data1$szcmttl == 10 | js_data1$szcmttl == 11 | js_data1$szcmttl == 12 | js_data1$szcmttl == 13 | js_data1$szcmttl == 14 | js_data1$szcmttl == 15 | js_data1$szcmttl == 16 | js_data1$szcmttl == 17 | js_data1$szcmttl == 18 | js_data1$szcmttl == 19] <- 1
#Combine 20-29 minutes
js_data1$szcmttl[js_data1$szcmttl == 20 | js_data1$szcmttl == 21 | js_data1$szcmttl == 22 | js_data1$szcmttl == 23 | js_data1$szcmttl == 24 | js_data1$szcmttl == 25 | js_data1$szcmttl == 26 | js_data1$szcmttl == 27 | js_data1$szcmttl == 28 | js_data1$szcmttl == 29] <- 2
#Combine 30-39 minutes
js_data1$szcmttl[js_data1$szcmttl == 30 | js_data1$szcmttl == 31 | js_data1$szcmttl == 32 | js_data1$szcmttl == 33 | js_data1$szcmttl == 34 | js_data1$szcmttl == 35 | js_data1$szcmttl == 36 | js_data1$szcmttl == 37 | js_data1$szcmttl == 38 | js_data1$szcmttl == 39] <- 3
#Combine 40-49 minutes
js_data1$szcmttl[js_data1$szcmttl == 40 | js_data1$szcmttl == 41 | js_data1$szcmttl == 42 | js_data1$szcmttl == 43 | js_data1$szcmttl
```

```r
== 44 | js_data1$szcmttl == 45 | js_data1$szcmttl == 46 | js_data1$szcmttl == 47 | js_data1$szcmttl == 48 | js_data1$szcmttl == 49] <- 4
#Combine 50-59 minutes
js_data1$szcmttl[js_data1$szcmttl == 50 | js_data1$szcmttl == 51 | js_data1$szcmttl == 52 | js_data1$szcmttl == 53 | js_data1$szcmttl == 54 | js_data1$szcmttl == 55 | js_data1$szcmttl == 56 | js_data1$szcmttl == 57 | js_data1$szcmttl == 58 | js_data1$szcmttl == 59] <- 5
#Combine 60-59 minutes
js_data1$szcmttl[js_data1$szcmttl == 60 | js_data1$szcmttl == 61 | js_data1$szcmttl == 62 | js_data1$szcmttl == 63 | js_data1$szcmttl == 64 | js_data1$szcmttl == 65 | js_data1$szcmttl == 66 | js_data1$szcmttl == 67 | js_data1$szcmttl == 68 | js_data1$szcmttl == 69] <- 6
#Combine 70-79 minutes
js_data1$szcmttl[js_data1$szcmttl == 70 | js_data1$szcmttl == 71 | js_data1$szcmttl == 72 | js_data1$szcmttl == 73 | js_data1$szcmttl == 74 | js_data1$szcmttl == 75 | js_data1$szcmttl == 76 | js_data1$szcmttl == 77 | js_data1$szcmttl == 78 | js_data1$szcmttl == 79] <- 7
#Combine 80-89 minutes
js_data1$szcmttl[js_data1$szcmttl == 80 | js_data1$szcmttl == 81 | js_data1$szcmttl == 82 | js_data1$szcmttl == 83 | js_data1$szcmttl == 84 | js_data1$szcmttl == 85 | js_data1$szcmttl == 86 | js_data1$szcmttl == 87 | js_data1$szcmttl == 88 | js_data1$szcmttl == 89] <- 8
#Combine remaining values to =<90 minutes
js_data1$szcmttl[js_data1$szcmttl >= 90] <- 9
table(js_data1$szcmttl)

#Work time
js_data1 <- subset(js_data1, js_data1$xjobhwk != 888 & js_data1$xjobhwk != 999)
table(js_data1$xjobhwk)
hist(js_data1$xjobhwk)
js_data1$xjobhwkLog <- log(js_data1$xjobhwk)
hist(js_data1$xjobhwkLog)

#Type of employement
table(js_data1$tpjob)
#Exclude missing values
js_data1 <- subset(js_data1, js_data1$tpjob != 7 & js_data1$tpjob != 8 & js_data1$tpjob != 9)
#Create dummy variables
js_data1$executive <- js_data1$tpjob
js_data1$executive [ js_data1$executive >= 2 ] <- 0
table(js_data1$executive)
js_data1$temporary <- js_data1$tpjob
js_data1$temporary [ js_data1$temporary >= 4 | js_data1$temporary <= 2 ] <- 2
table(js_data1$temporary)
js_data1$dispatched <- js_data1$tpjob
```

```r
table(js_data1$dispatched)
js_data1$dispatched [ js_data1$dispatched >= 5 | js_data1$dispatched <= 3 ] <- 3
js_data1$self_employed <- js_data1$tpjob
table(js_data1$self_employed)
js_data1$self_employed [ js_data1$self_employed >= 6 | js_data1$self_employed <= 4 ] <- 4
js_data1$family_worker  <- js_data1$tpjob
js_data1$family_worker [ js_data1$family_worker <= 5] <- 5
table(js_data1$family_worker)

#Goverment agency employee
js_data1$GovermentAgency <- js_data1$szttlsta
table(js_data1$GovermentAgency)
#Moving goverment agency employees to highest category (100)
js_data1$GovermentAgency[js_data1$GovermentAgency == 12] <- 100
#Everyhing else to 0
js_data1$GovermentAgency[js_data1$GovermentAgency <= 99] <- 0
#Moving goverment agency employees to 1
js_data1$GovermentAgency[js_data1$GovermentAgency == 100] <- 1
#Marking GovernmentAgency as categorical variable
js_data1$GovermentAgencyFactor <- factor(js_data1$GovermentAgency, labels=c("Normal","GovernmentAgency"))
table(js_data1$GovermentAgency)

#Health (satisfaction with health and physical condition)
table(js_data1$st5hlthy)
#Removing na's
js_data1 <- subset(js_data1, js_data1$st5hlthy != 9)
#Reversing order for easier understanding
js_data1$st5hlthy[js_data1$st5hlthy == 5] <- 0
js_data1$st5hlthy[js_data1$st5hlthy == 1] <- 5
js_data1$st5hlthy[js_data1$st5hlthy == 4] <- 1
js_data1$st5hlthy[js_data1$st5hlthy == 2] <- 4
js_data1$st5hlthy[js_data1$st5hlthy == 3] <- 2
js_data1$st5hlthy[js_data1$st5hlthy == 4] <- 3
js_data1$st5hlthy[js_data1$st5hlthy == 5] <- 4
table(js_data1$st5hlthy)

#Satisfaction with living place
table(js_data1$st5areay)
#Removing na's
js_data1 <- subset(js_data1, js_data1$st5areay != 9)
#Reversing order for easier understanding
js_data1$st5areay[js_data1$st5areay == 5] <- 0
js_data1$st5areay[js_data1$st5areay == 1] <- 5
js_data1$st5areay[js_data1$st5areay == 4] <- 1
js_data1$st5areay[js_data1$st5areay == 2] <- 4
js_data1$st5areay[js_data1$st5areay == 3] <- 2
js_data1$st5areay[js_data1$st5areay == 4] <- 3
js_data1$st5areay[js_data1$st5areay == 5] <- 4
table(js_data1$st5areay)
```

```
#Leisure satisfaction
table(js_data1$st5leisy)
#Removing na's
js_data1 <- subset(js_data1, js_data1$st5leisy != 9)
#Reversing order for easier understanding
js_data1$st5leisy[js_data1$st5leisy == 5] <- 0
js_data1$st5leisy[js_data1$st5leisy == 1] <- 5
js_data1$st5leisy[js_data1$st5leisy == 4] <- 1
js_data1$st5leisy[js_data1$st5leisy == 2] <- 4
js_data1$st5leisy[js_data1$st5leisy == 3] <- 2
js_data1$st5leisy[js_data1$st5leisy == 4] <- 3
js_data1$st5leisy[js_data1$st5leisy == 5] <- 4
table(js_data1$st5leisy)

#Inclusion of different leisure activities
#Television
table(js_data1$hrtv)
#Removing na's
js_data1 <- subset(js_data1, js_data1le$hrtv != 999)
#Combining everything above 7 in one category
js_data1$hrtv[js_data1$hrtv >= 7] <- 7

#Number of books (comics,magazines) per month
table(js_data1$fq5read)
#Removing na's
js_data1 <- subset(js_data1, js_data1$fq5read != 9)

#Playing sports
table(js_data1$fqsport)
#Removing na's
js_data1 <- subset(js_data1, js_data1$fqsport != 9)
#Changing order for easier understanding
js_data1$fqsport[js_data1$fqsport == 5] <- 0
js_data1$fqsport[js_data1$fqsport == 1] <- 5
js_data1$fqsport[js_data1$fqsport == 4] <- 1
js_data1$fqsport[js_data1$fqsport == 2] <- 4
js_data1$fqsport[js_data1$fqsport == 3] <- 2
js_data1$fqsport[js_data1$fqsport == 4] <- 3
js_data1$fqsport[js_data1$fqsport == 5] <- 4
table(js_data1$fqsport)

#Meeting friends (frequency)
table(js_data1$fq7frsee)
#Removing na's
js_data1 <- subset(js_data1, js_data1$fq7frsee != 9)
#Changing order for better understanding + combining 1(every day), 2(several times a week) to "several times a week"
js_data1$fq7frsee[js_data1$fq7frsee == 1 ] <- 2
#Change order for better undrstanding
js_data1$fq7frsee[js_data1$fq7frsee == 7] <- 0
js_data1$fq7frsee[js_data1$fq7frsee == 6] <- 1
js_data1$fq7frsee[js_data1$fq7frsee == 2] <- 6
js_data1$fq7frsee[js_data1$fq7frsee == 5] <- 2
```

```
js_data1$fq7frsee[js_data1$fq7frsee == 3] <- 5
js_data1$fq7frsee[js_data1$fq7frsee == 4] <- 3
js_data1$fq7frsee[js_data1$fq7frsee == 5] <- 4
js_data1$fq7frsee[js_data1$fq7frsee == 6] <- 5
table(js_data1$fq7frsee)

#Civic group member
#Combining membership in citizens movement/consumers cooperative groups and membership in social service group
table(js_data1$memcivil)
js_data1$civic <- js_data1$memcivil
table(js_data1$civic)
js_data1$civic[js_data1$civic == 2] <- 0
#Combining membership and volunteer organisation and membership in civil society
js_data1$civic[js_data1$memvlntr == 1] <- 1
#Removing na's
js_data1 <- subset(js_data1, js_data1$civic != 9)
table(js_data1$civic)

#Hobbygroup member
table(js_data1$memhobby)
js_data1 <- subset(js_data1, js_data1$memhobby != 9)
js_data1$memhobby[js_data1$memhobby == 2] <- 0
table(js_data1$memhobby)

#Barplot for job satifaction distribution
barplot(table(js_data1$st5job), col="black", names.arg = c("satisfied (1)", "somewhat satisfied (2)", "neither (3)","somewhat dissatisfied (4)", "dissatisfied (5)"), cex.axis=0.8, cex.names=0.7, xlab= "Level of job dissatisfaction", ylab="Number of responses", main=NULL)

#Combining last categories of job satisfaction and changing order for better understanding
js_data1$st5job[js_data1$st5job == 5 | js_data1$st5job == 4 ] <- 0
#Changing order for better understanding
js_data1$st5job[js_data1$st5job == 1 ] <- 4
js_data1$st5job[js_data1$st5job == 3 ] <- 1
js_data1$st5job[js_data1$st5job == 4 ] <- 3
table(js_data1$st5job)

#Barplot for new distribution of job satifaction
barplot(table(js_data1$st5job), col="black", names.arg = c( "(somewhat) dissatisfied (0)", "neither (1)", "somewhat satisfied (2)", "satisfied (3)"), cex.axis=0.8, cex.names=0.8, xlab= "Level of job satisfaction", ylab="Number of responses", main=NULL)

#Descriptive Statistics
summary(js_data1$st5job)
se(js_data1$st5job)
summary(js_data1$ageb)
se(js_data1$ageb)
```

```
summary(js_data1$male)
se(js_data1$male)
summary(js_data1$never_married)
se(js_data1$never_married)
summary(js_data1$divorced)
se(js_data1$divorced)
summary(js_data1$widowed)
se(js_data1$widowed)
summary(js_data1$ccnumttl)
se(js_data1$ccnumttl)
summary(js_data1$graduatedCollege)
se(js_data1$graduatedCollege)
summary(js_data1$szincomx)
se(js_data1$szincomx)
summary(js_data1$xjobhwk)
se(js_data1$xjobhwk)
summary(js_data1$executive)
se(js_data1$executive)
summary(js_data1$temporary)
se(js_data1$temporary)
summary(js_data1$dispatched)
se(js_data1$dispatched)
summary(js_data1$self_employed)
se(js_data1$self_employed)
summary(js_data1$family_worker)
se(js_data1$family_worker)
summary(js_data1$GovermentAgency)
se(js_data1$GovermentAgency)
summary(js_data1$szcmttl)
se(js_data1$szcmttl)
summary(js_data1$st5hlthy)
se(js_data1$st5hlthy)
summary(js_data1$st5areay)
se(js_data1$st5areay)
summary(js_data1$st5leisy)
se(js_data1$st5leisy)
summary(js_data1$civic)
se(js_data1$civic)
summary(js_data1$memhobby)
se(js_data1$memhobby)
summary(js_data1$fq7frsee)
se(js_data1$fq7frsee)
summary(js_data1$fqsport)
se(js_data1$fqsport)
summary(js_data1$fq5read)
se(js_data1$fq5read)
summary(js_data1$hrtv)
se(js_data1$hrtv)


#Extimating regression for main sample
ols1 <- lm(st5job ~ ageb + agebSquare + male + never_married +
    divorced + widowed + ccnumttl + graduatedCollege + szincomx +
```

```
    xjobhwkLog + executive + temporary + dispatched + self_employed +
     family_worker + GovermentAgency + szcmttl + st5hlthy + st5areay+
     st5leisy + civic + memhobby + fq7frsee + fqsport + fq5read +
    hrtv,  data=js_data1)
summary(ols1)

#Breusch-Pegan test
bptest(ols1)
#Result:BP = BP = 47.7937, df = 26, p-value = 0.005707
#Estimation of model with heteroskedasticity robust standard errors
coeftest(ols1, df = Inf, vcov = vcovHC(ols1, type="HC"))

#Test for multicollianrity
vif(ols1)

#Correlation Matrix
attach(js_data1)
d <- data.frame(ageb, agebSquare, male, never_married, divorced,
    widowed, ccnumttl, graduatedCollege, szincomx, xjobhwkLog,
    executive, temporary, dispatched, self_employed, family_worker,
    GovermentAgency, szcmttl, st5hlthy, st5areay, st5leisy, civic,
    memhobby, fq7frsee, fqsport, fq5read, hrtv)
cor(d) # get correlations
as.matrix(cor(d))
detach(js_data1)

#Estimating Model acounding for hetereskedasticity using sandwitch
    estimator
(ordered_logit1 <- lrm(st5job ~ ageb + agebSquare + male + never_
    married + divorced + widowed + ccnumttl + graduatedCollege +
    szincomx + xjobhwkLog + executive + temporary + dispatched + self
    _employed + family_worker + GovermentAgency + szcmttl + st5hlthy
    + st5areay+ st5leisy + civic + memhobby + fq7frsee + fqsport +
    fq5read + hrtv, x=TRUE, y=TRUE, data=js_data1))
bptest(ordered_logit1)
#Result: BP = 46.2512, df = 25, p-value = 0.006001
(g <- robcov(ordered_logit1))

#Calculate McFadden Pseudo-R2
vglm1 <- vglm(st5job ~ 1, family=propodds, data=js_data1)
LLf   <- logLik(g)
LL0   <- logLik(vglm1)
as.vector(1 - (LLf / LL0))

#Calculate Loglikelihood
(logLik(g))

#VIFs for multicolianarity
vif(g)


#Estimation of regression for main famale sample
js_data1Fem <- js_data1
```

```
#Ereasing male sample
js_data1Fem <- subset(js_data1Fem, js_data1Fem$sexa != 1)
#Estimation OLS
ols1Fem <- lm(st5job ~ ageb + agebSquare + never_married + divorced +
     widowed + ccnumttl +  graduatedCollege + szincomx + xjobhwkLog +
     executive + temporary + dispatched + self_employed + family_
    worker + GovermentAgency + szcmttl + st5hlthy + st5areay+
    st5leisy + civic + memhobby + fq7frsee + fqsport + fq5read + hrtv
     , data=js_data1Fem)
summary(ols1Fem)

#Descriptive Statistics
summary(js_data1Fem$st5job)
se(js_data1Fem$st5job)
summary(js_data1Fem$ageb)
se(js_data1Fem$ageb)
summary(js_data1Fem$never_married)
se(js_data1Fem$never_married)
summary(js_data1Fem$divorced)
se(js_data1Fem$divorced)
summary(js_data1Fem$widowed)
se(js_data1Fem$widowed)
summary(js_data1Fem$ccnumttl)
se(js_data1Fem$ccnumttl)
summary(js_data1Fem$graduatedCollege)
se(js_data1Fem$graduatedCollege)
summary(js_data1Fem$szincomx)
se(js_data1Fem$szincomx)
summary(js_data1Fem$xjobhwk)
se(js_data1Fem$xjobhwk)
summary(js_data1Fem$executive)
se(js_data1Fem$executive)
summary(js_data1Fem$temporary)
se(js_data1Fem$temporary)
summary(js_data1Fem$dispatched)
se(js_data1Fem$dispatched)
summary(js_data1Fem$self_employed)
se(js_data1Fem$self_employed)
summary(js_data1Fem$family_worker)
se(js_data1Fem$family_worker)
summary(js_data1Fem$GovermentAgency)
se(js_data1Fem$GovermentAgency)
summary(js_data1Fem$szcmttl)
se(js_data1Fem$szcmttl)
summary(js_data1Fem$st5hlthy)
se(js_data1Fem$st5hlthy)
summary(js_data1Fem$st5areay)
se(js_data1Fem$st5areay)
summary(js_data1Fem$st5leisy)
se(js_data1Fem$st5leisy)
summary(js_data1Fem$civic)
se(js_data1Fem$civic)
summary(js_data1Fem$memhobby)
```

```
se(js_data1Fem$memhobby)
summary(js_data1Fem$fq7frsee)
se(js_data1Fem$fq7frsee)
summary(js_data1Fem$fqsport)
se(js_data1Fem$fqsport)
summary(js_data1Fem$fq5read)
se(js_data1Fem$fq5read)
summary(js_data1Fem$hrtv)
se(js_data1Fem$hrtv)

#Breusch-Pagan test
bptest(ols1Fem)
#Results: BP = 53.3928, df = 25, p-value = 0.000796
#Estimation of model with heteroskedasticity robust standard errors
coeftest(ols1Fem, df = Inf, vcov = vcovHC(ols1Fem, type="HC"))

#Test for multicollianrity
vif(ols1Fem)

#Correlation Matrix
attach(js_data1Fem)
ols1Femcor <- data.frame(ageb, agebSquare, never_married, divorced,
    widowed, ccnumttl, graduatedCollege, szincomx, xjobhwkLog,
    executive, temporary, dispatched, self_employed, family_worker,
    GovermentAgency, szcmttl, st5hlthy, st5areay, st5leisy, civic,
    memhobby, fq7frsee, fqsport, fq5read, hrtv)
cor(ols1Femcor) # get correlations
as.matrix(cor(ols1Femcor))
detach(js_data1Fem)

#Extimation ordered logit female sample
(ordered_logit1Fem <- lrm(st5job ~ ageb + agebSquare + never_married
    + divorced + widowed + ccnumttl + graduatedCollege + szincomx +
    xjobhwkLog + executive + temporary + dispatched + self_employed +
     family_worker + GovermentAgency + szcmttl + st5hlthy + st5areay+
     st5leisy + civic + memhobby + fq7frsee + fqsport + fq5read +
    hrtv, x=TRUE, y=TRUE, data=js_data1Fem))

#Estimating Model acounding for hetereskedasticity using sandwitch
    estimator
bptest(ordered_logit1Fem)
#Results: BP = 55.975, df = 24, p-value = 0.000231
(f <- robcov(ordered_logit1Fem))

#Calculate McFadden Pseudo-R2
vglm2 <- vglm(st5job ~ 1, family=propodds, data=js_data1Fem)
LLf   <- logLik(f)
LL0   <- logLik(vglm2)
as.vector(1 - (LLf / LL0))

#Calculate Loglikelihood
(logLik(f))
```

```
#VIFs for multicolianarity check
vif(f)


#Estmation of regression for male sample
js_data1Mal <- js_data1
#Ereasing female sample
js_data1Mal <- subset(js_data1Mal, js_data1Mal$sexa != 0)
#Estimation OLS
ols1Mal <- lm(st5job ~ ageb + agebSquare + never_married + divorced +
    widowed + ccnumttl + graduatedCollege + szincomx + xjobhwkLog +
    executive + temporary + dispatched + self_employed + family_
    worker + GovermentAgency + szcmttl + st5hlthy + st5areay +
    st5leisy + civic + memhobby + fq7frsee + fqsport + fq5read + hrtv
    , data=js_data1Mal)
summary(ols1Mal)

#Descriptive Statistics
summary(js_data1Mal$st5job)
se(js_data1Mal$st5job)
summary(js_data1Mal$ageb)
se(js_data1Mal$ageb)
summary(js_data1Mal$male)
se(js_data1Mal$male)
summary(js_data1Mal$never_married)
se(js_data1Mal$never_married)
summary(js_data1Mal$divorced)
se(js_data1Mal$divorced)
summary(js_data1Mal$widowed)
se(js_data1Mal$widowed)
summary(js_data1Mal$ccnumttl)
se(js_data1Mal$ccnumttl)
summary(js_data1Mal$graduatedCollege)
se(js_data1Mal$graduatedCollege)
summary(js_data1Mal$szincomx)
se(js_data1Mal$szincomx)
summary(js_data1Mal$xjobhwk)
se(js_data1Mal$xjobhwk)
summary(js_data1Mal$executive)
se(js_data1Mal$executive)
summary(js_data1Mal$temporary)
se(js_data1Mal$temporary)
summary(js_data1Mal$dispatched)
se(js_data1Mal$dispatched)
summary(js_data1Mal$self_employed)
se(js_data1Mal$self_employed)
summary(js_data1Mal$family_worker)
se(js_data1Mal$family_worker)
summary(js_data1Mal$GovermentAgency)
se(js_data1Mal$GovermentAgency)
summary(js_data1Mal$szcmttl)
se(js_data1Mal$szcmttl)
summary(js_data1Mal$st5hlthy)
```

```
se(js_data1Mal$st5hlthy)
summary(js_data1Mal$st5areay)
se(js_data1Mal$st5areay)
summary(js_data1Mal$st5leisy)
se(js_data1Mal$st5leisy)
summary(js_data1Mal$civic)
se(js_data1Mal$civic)
summary(js_data1Mal$memhobby)
se(js_data1Mal$memhobby)
summary(js_data1Mal$fq7frsee)
se(js_data1Mal$fq7frsee)
summary(js_data1Mal$fqsport)
se(js_data1Mal$fqsport)
summary(js_data1Mal$fq5read)
se(js_data1Mal$fq5read)
summary(js_data1Mal$hrtv)
se(js_data1Mal$hrtv)

#Breusch-Pagan test
bptest(ols1Mal)
# #Results: BP = 22.2914, df = 25, p-value = 0.6189, therefore no
    robust SEs need to be estimated
#Estimation of model with heteroskedasticity robust standard errors
coeftest(ols1Mal, df = Inf, vcov = vcovHC(ols1Mal, type="HC"))

#Test for multicollianrity
vif(ols1Mal)

#Correlation Matrix
attach(js_data1Mal)
ols1Malcor <- data.frame(ageb, agebSquare, never_married, divorced,
    widowed, ccnumttl, graduatedCollege, szincomx, xjobhwkLog,
    executive, temporary, dispatched, self_employed, family_worker,
    GovermentAgency, szcmttl, st5hlthy, st5areay,  st5leisy,  civic,
    memhobby, fq7frsee, fqsport, fq5read, hrtv)
cor(ols1Malcor) # get correlations
as.matrix(cor(ols1Malcor))
detach(js_data1Mal)

#Extimation ordered logit
(ordered_logit1Mal<- lrm(st5job ~ ageb + agebSquare + never_married +
     divorced + widowed + ccnumttl + graduatedCollege + szincomx +
    xjobhwkLog + executive + temporary + dispatched + self_employed +
     family_worker + GovermentAgency + szcmttl + st5hlthy + st5areay+
     st5leisy + civic + memhobby + fq7frsee + fqsport + fq5read +
    hrtv, x=TRUE, y=TRUE, data=js_data1Mal))

#Estimating Model acounding for hetereskedasticity using sandwitch
    estimator
bptest(ordered_logit1Mal)
#Results:BP = 22.2914, df = 25, p-value = 0.6189
(m <- robcov(ordered_logit1Mal))
```

```
#VIFs for multicolianarity check
vif(m)

#Calculate McFadden Pseudo-R2
vglm3 <- vglm(st5job ~ 1, family=propodds, data=js_data1Mal)
LLf   <- logLik(m)
LL0   <- logLik(vglm3)
as.vector(1 - (LLf / LL0))

#Calculate Loglikelihood
(logLik(m))

#Request OLS Main
stargazer(list(coeftest(ols1, df = Inf, vcov = vcovHC(ols1, type="HC")), coeftest(ols1Fem, df = Inf, vcov = vcovHC(ols1Fem, type="HC")),coeftest(ols1Mal, df = Inf, vcov = vcovHC(ols1Mal, type="HC"))))
#Request Ordered Logit Main
stargazer(list(g, f, m))


#Models without irrelvant variables
summary(ols1)
coeftest(ols1, df = Inf, vcov = vcovHC(ols1, type="HC"))
#Dummy variable self-employed least significant (might be to small fraction (2025/184))
ols1a <- lm(st5job ~ ageb + agebSquare + male + never_married + divorced + widowed + ccnumttl + graduatedCollege + szincomx + xjobhwkLog + executive + temporary + dispatched + family_worker + GovermentAgency + szcmttl + st5hlthy + st5areay+ st5leisy + civic + memhobby + fq7frsee + fqsport + fq5read + hrtv,  data=js_data1)
summary(ols1a)
#Adjusted Rquare higher
coeftest(ols1a, df = Inf, vcov = vcovHC(ols1a, type="HC"))
#Adjusted Rquare higher
#Dummy variable temporary now least significant
ols1b <- lm(st5job ~ ageb + agebSquare + male + never_married + divorced + widowed + ccnumttl + graduatedCollege + szincomx + xjobhwkLog + executive + dispatched + family_worker + GovermentAgency + szcmttl + st5hlthy + st5areay+ st5leisy + civic + memhobby + fq7frsee + fqsport + fq5read + hrtv,  data=js_data1)
summary(ols1b)
coeftest(ols1b, df = Inf, vcov = vcovHC(ols1b, type="HC"))
#Adjusted Rsquare higher
#Number of children now least significant
ols1c <- lm(st5job ~ ageb + agebSquare + male + never_married + divorced + widowed + graduatedCollege + szincomx + xjobhwkLog + executive + dispatched + family_worker + GovermentAgency + szcmttl + st5hlthy + st5areay+ st5leisy + civic + memhobby + fq7frsee + fqsport + fq5read + hrtv,  data=js_data1)
summary(ols1c)
```

```
coeftest(ols1c, df = Inf, vcov = vcovHC(ols1c, type="HC"))
#Adjusted Rsquare higher
#Commuting time now least significant
ols1d <- lm(st5job ~ ageb + agebSquare + male + never_married +
    divorced + widowed + graduatedCollege + szincomx + xjobhwkLog +
    executive + dispatched + family_worker + GovermentAgency +
    st5hlthy + st5areay+ st5leisy + civic + memhobby + fq7frsee +
    fqsport + fq5read + hrtv, data=js_data1)
summary(ols1d)
coeftest(ols1d, df = Inf, vcov = vcovHC(ols1d, type="HC"))
#Adjusted Rsquare higher
#Being divorced least signifcant (due to the small number)
ols1e <- lm(st5job ~ ageb + agebSquare + male + never_married +
    widowed + graduatedCollege + szincomx + xjobhwkLog + executive +
    dispatched + family_worker + GovermentAgency + st5hlthy +
    st5areay+ st5leisy + civic + memhobby + fq7frsee + fqsport +
    fq5read + hrtv, data=js_data1)
summary(ols1e)
coeftest(ols1e, df = Inf, vcov = vcovHC(ols1e, type="HC"))
#Adjusted Rsquare higher
#Being family worker now least signficant
ols1f <- lm(st5job ~ ageb + agebSquare + male + never_married +
    widowed + graduatedCollege + szincomx + xjobhwkLog + executive +
     dispatched + GovermentAgency + st5hlthy + st5areay+ st5leisy +
    civic + memhobby + fq7frsee + fqsport + fq5read + hrtv, data=js_
    data1)
summary(ols1f)
coeftest(ols1f, df = Inf, vcov = vcovHC(ols1f, type="HC"))
#Adjusted Rsquare higher
#Being executive least signficant
ols1g <- lm(st5job ~ ageb + agebSquare + male + never_married +
    widowed + graduatedCollege + szincomx + xjobhwkLog + dispatched
     + GovermentAgency + st5hlthy + st5areay+ st5leisy + civic +
    memhobby + fq7frsee + fqsport + fq5read + hrtv, data=js_data1)
summary(ols1g)
coeftest(ols1g, df = Inf, vcov = vcovHC(ols1g, type="HC"))
#Adjusted Rsquare higher
#Seeing of friends now least significant
ols1h <- lm(st5job ~ ageb + agebSquare + male + never_married +
    widowed + graduatedCollege + szincomx + xjobhwkLog + dispatched
     + GovermentAgency + st5hlthy + st5areay+ st5leisy + civic +
    memhobby + fqsport + fq5read + hrtv, data=js_data1)
summary(ols1h)
coeftest(ols1h, df = Inf, vcov = vcovHC(ols1h, type="HC"))
#Hours of watching TV now least signficant
ols1i <- lm(st5job ~ ageb + agebSquare + male + never_married +
    widowed + graduatedCollege + szincomx + xjobhwkLog + dispatched
     + GovermentAgency + st5hlthy + st5areay+ st5leisy + civic +
    memhobby + fqsport + fq5read , data=js_data1)
summary(ols1i)
coeftest(ols1i, df = Inf, vcov = vcovHC(ols1i, type="HC"))
#Adjusted Rsquare higher
#Deing widowed now least significant
```

```
ols1j <- lm(st5job ~ ageb + agebSquare + male + never_married +
    graduatedCollege + szincomx + xjobhwkLog   + dispatched +
    GovermentAgency  + st5hlthy + st5areay+ st5leisy + civic +
    memhobby  + fqsport + fq5read  ,  data=js_data1)
summary(ols1j)
coeftest(ols1j, df = Inf, vcov = vcovHC(ols1j, type="HC"))
#Adjusted Rsquare higher
#Frequency of sports now least signficant
ols1k <- lm(st5job ~ ageb + agebSquare + male + never_married +
    graduatedCollege + szincomx + xjobhwkLog + dispatched +
    GovermentAgency  + st5hlthy + st5areay+ st5leisy + civic +
    memhobby  + fq5read  ,  data=js_data1)
summary(ols1k)
coeftest(ols1k, df = Inf, vcov = vcovHC(ols1k, type="HC"))
#Adjusted Rsquare the same as before - stopping elimination of
    irrelvant values here because adjusted r square will decrease
    with next least significant variable (checked)


#Models without irrelvant variables (female sample)
summary(ols1Fem)
coeftest(ols1Fem, df = Inf, vcov = vcovHC(ols1Fem, type="HC"))
#Number of children least signifcant
ols1Fema <- lm(st5job ~ ageb + agebSquare  + never_married + divorced
     + widowed  + graduatedCollege + szincomx + xjobhwkLog +
    executive + temporary + dispatched + self_employed + family_
    worker + GovermentAgency + szcmttl + st5hlthy + st5areay+
    st5leisy + civic + memhobby + fq7frsee + fqsport + fq5read + hrtv
    ,  data=js_data1Fem)
summary(ols1Fema)
#Adjusted Rsquare higher
coeftest(ols1Fema, df = Inf, vcov = vcovHC(ols1Fema, type="HC"))
#self-eployment now least significant
ols1Femb <- lm(st5job ~ ageb + agebSquare + never_married + divorced
    + widowed  + graduatedCollege + szincomx + xjobhwkLog + executive
     + temporary + dispatched  + family_worker + GovermentAgency +
    szcmttl + st5hlthy + st5areay+ st5leisy + civic + memhobby +
    fq7frsee + fqsport + fq5read + hrtv,  data=js_data1Fem)
summary(ols1Femb)
#Adjusted Rsquare higher
coeftest(ols1Femb, df = Inf, vcov = vcovHC(ols1Femb, type="HC"))
#Seeing friends now least significant
ols1Femc <- lm(st5job ~ ageb + agebSquare + never_married + divorced
    + widowed  + graduatedCollege + szincomx + xjobhwkLog + executive
     + temporary + dispatched  + family_worker + GovermentAgency +
    szcmttl + st5hlthy + st5areay+ st5leisy + civic + memhobby  +
    fqsport + fq5read + hrtv,  data=js_data1Fem)
summary(ols1Femc)
#Adjusted Rsquare higher
coeftest(ols1Femc, df = Inf, vcov = vcovHC(ols1Femc, type="HC"))
#doing sports now now least significant
ols1Femd <- lm(st5job ~ ageb + agebSquare + never_married + divorced
    + widowed  + graduatedCollege + szincomx + xjobhwkLog + executive
```

```
 + temporary + dispatched  + family_worker + GovermentAgency +
 szcmttl + st5hlthy + st5areay+ st5leisy + civic + memhobby  +
 fq5read + hrtv ,  data=js_data1Fem)
summary(ols1Femd)
#Adjusted Rsquare higher
coeftest(ols1Femd, df = Inf, vcov = vcovHC(ols1Femd, type="HC"))
#watching TV now least significant
ols1Feme <- lm(st5job ~ ageb + agebSquare + never_married + divorced
 + widowed  + graduatedCollege + szincomx + xjobhwkLog + executive
 + temporary + dispatched  + family_worker + GovermentAgency +
 szcmttl + st5hlthy + st5areay+ st5leisy + civic + memhobby  +
 fq5read ,  data=js_data1Fem)
summary(ols1Feme)
#Adjusted Rsquare higher
coeftest(ols1Feme, df = Inf, vcov = vcovHC(ols1Feme, type="HC"))
#Beeing widowed now least significant
ols1Femf <- lm(st5job ~ ageb + agebSquare + never_married + divorced
 + graduatedCollege + szincomx + xjobhwkLog + executive +
 temporary + dispatched  + family_worker + GovermentAgency +
 szcmttl + st5hlthy + st5areay+ st5leisy + civic + memhobby  +
 fq5read ,  data=js_data1Fem)
summary(ols1Femf)
#Adjusted Rsquare higher
coeftest(ols1Femf, df = Inf, vcov = vcovHC(ols1Femf, type="HC"))
#executive postion now least significant
ols1Femg <- lm(st5job ~ ageb + agebSquare + never_married + divorced
 + graduatedCollege + szincomx + xjobhwkLog  + temporary +
 dispatched  + family_worker + GovermentAgency + szcmttl +
 st5hlthy + st5areay+ st5leisy + civic + memhobby + fq5read ,
 data=js_data1Fem)
summary(ols1Femg)
#Adjusted Rsquare higher
coeftest(ols1Femg, df = Inf, vcov = vcovHC(ols1Femg, type="HC"))
#Position in governmental agency now least significant
ols1Femh <- lm(st5job ~ ageb + agebSquare + never_married + divorced
 + graduatedCollege + szincomx + xjobhwkLog  + temporary +
 dispatched  + family_worker  + szcmttl + st5hlthy + st5areay+
 st5leisy + civic + memhobby + fq5read ,  data=js_data1Fem)
summary(ols1Femh)
#Adjusted Rsquare higher
coeftest(ols1Femh, df = Inf, vcov = vcovHC(ols1Femh, type="HC"))
#Being family worker now least significant
ols1Feml <- lm(st5job ~ ageb + agebSquare + never_married + divorced
 + graduatedCollege + szincomx + xjobhwkLog  + temporary +
 dispatched  + szcmttl + st5hlthy + st5areay+ st5leisy + civic +
 memhobby + fq5read ,  data=js_data1Fem)
summary(ols1Feml)
#Adjusted Rsquare the lower than before - stopping elimination of
 irrelvant values here with previous model (ols1Femh)
coeftest(ols1Feml, df = Inf, vcov = vcovHC(ols1Feml, type="HC"))


#Models without irrelvant variables (male sample)
```

```r
summary(ols1Mal)
coeftest(ols1Mal, df = Inf, vcov = vcovHC(ols1Mal, type="HC"))
#self_employed least significant
ols1Mala <- lm(st5job ~ ageb + agebSquare + never_married + divorced
    + widowed + ccnumttl + graduatedCollege + szincomx + xjobhwkLog +
     executive + temporary + dispatched  + family_worker +
    GovermentAgency + szcmttl + st5hlthy + st5areay+ st5leisy + civic
     + memhobby + fq7frsee + fqsport + fq5read + hrtv, js_data1Mal)
summary(ols1Mala)
#Adjusted Rsquare higher
coeftest(ols1Mala, df = Inf, vcov = vcovHC(ols1Mala, type="HC"))
#Seein friends now least significant
ols1Malb <- lm(st5job ~ ageb + agebSquare + never_married + divorced
    + widowed + ccnumttl + graduatedCollege + szincomx + xjobhwkLog +
     executive + temporary + dispatched  + family_worker +
    GovermentAgency + szcmttl + st5hlthy + st5areay+ st5leisy + civic
     + memhobby + fqsport + fq5read + hrtv, js_data1Mal)
summary(ols1Malb)
#Adjusted Rsquare higher
coeftest(ols1Malb, df = Inf, vcov = vcovHC(ols1Malb, type="HC"))
#Watching TV now least significant
ols1Malc <- lm(st5job ~ ageb + agebSquare + never_married + divorced
    + widowed + ccnumttl + graduatedCollege + szincomx + xjobhwkLog +
     executive + temporary + dispatched  + family_worker +
    GovermentAgency + szcmttl + st5hlthy + st5areay+ st5leisy + civic
     + memhobby + fqsport + fq5read, js_data1Mal)
summary(ols1Malc)
#Adjusted Rsquare higher
coeftest(ols1Malc, df = Inf, vcov = vcovHC(ols1Malc, type="HC"))
#Number of children now least significant
ols1Mald <- lm(st5job ~ ageb + agebSquare + never_married + divorced
    + widowed  + graduatedCollege + szincomx + xjobhwkLog + executive
     + temporary + dispatched  + family_worker + GovermentAgency +
    szcmttl + st5hlthy + st5areay+ st5leisy + civic + memhobby +
    fqsport + fq5read, js_data1Mal)
summary(ols1Mald)
#Adjusted Rsquare higher
coeftest(ols1Mald, df = Inf, vcov = vcovHC(ols1Mald, type="HC"))
#Beeing tempory worker now least significant
ols1Male <- lm(st5job ~ ageb + agebSquare + never_married + divorced
    + widowed  + graduatedCollege + szincomx + xjobhwkLog + executive
      + dispatched  + family_worker + GovermentAgency + szcmttl +
    st5hlthy + st5areay+ st5leisy + civic + memhobby + fqsport +
    fq5read, js_data1Mal)
summary(ols1Male)
#Adjusted Rsquare higher
coeftest(ols1Male, df = Inf, vcov = vcovHC(ols1Male, type="HC"))
#Beeing executive now least significant
ols1Malf <- lm(st5job ~ ageb + agebSquare + never_married + divorced
    + widowed  + graduatedCollege + szincomx + xjobhwkLog +
    dispatched  + family_worker + GovermentAgency + szcmttl +
    st5hlthy + st5areay+ st5leisy + civic + memhobby + fqsport +
    fq5read, js_data1Mal)
```

```
summary(ols1Malf)
#Adjusted Rsquare higher
coeftest(ols1Malf, df = Inf, vcov = vcovHC(ols1Malf, type="HC"))
#Beeing divorced now least significant
ols1Malg <- lm(st5job ~ ageb + agebSquare + never_married + widowed
    + graduatedCollege + szincomx + xjobhwkLog + dispatched  + family
    _worker + GovermentAgency + szcmttl + st5hlthy + st5areay+
    st5leisy + civic + memhobby + fqsport + fq5read, js_data1Mal)
summary(ols1Malg)
#Adjusted Rsquare higher
coeftest(ols1Malg, df = Inf, vcov = vcovHC(ols1Malg, type="HC"))
#Commuting hours now least significant
ols1Malh <- lm(st5job ~ ageb + agebSquare + never_married + widowed
    + graduatedCollege + szincomx + xjobhwkLog + dispatched  + family
    _worker + GovermentAgency + st5hlthy + st5areay+ st5leisy + civic
     + memhobby + fqsport + fq5read, js_data1Mal)
summary(ols1Malh)
#Adjusted Rsquare higher
coeftest(ols1Malh, df = Inf, vcov = vcovHC(ols1Malh, type="HC"))
#Doing sport now least significant
ols1Mali <- lm(st5job ~ ageb + agebSquare + never_married + widowed
    + graduatedCollege + szincomx + xjobhwkLog + dispatched  + family
    _worker + GovermentAgency + st5hlthy + st5areay+ st5leisy + civic
     + memhobby + fq5read, js_data1Mal)
summary(ols1Mali)
#Adjusted Rsquare higher
coeftest(ols1Mali, df = Inf, vcov = vcovHC(ols1Mali, type="HC"))
#area of living now least significant
ols1Malj <- lm(st5job ~ ageb + agebSquare + never_married + widowed
    + graduatedCollege + szincomx + xjobhwkLog + dispatched  + family
    _worker + GovermentAgency + st5hlthy + st5leisy + civic +
    memhobby + fq5read, js_data1Mal)
summary(ols1Malj)
#Adjusted Rsquare the lower than before - stopping elimination of
    irrelvant values here with previous model (ols1Mali)
coeftest(ols1Mali, df = Inf, vcov = vcovHC(ols1Mali, type="HC"))

#Request Models without irrelevant varaibles
stargazer(list(coeftest(ols1j, df = Inf, vcov = vcovHC(ols1j, type="
    HC")), coeftest(ols1Femh, df = Inf, vcov = vcovHC(ols1Femh, type=
    "HC")), coeftest(ols1Mali, df = Inf, vcov = vcovHC(ols1Mali, type
    ="HC"))))

#VIFs for multicolianarity check
vif(ols1j)
vif(ols1Femh)
vif(ols1Mali)


#Models derived from Form A
js_data1a <- js_data1

#Adding new variables
```

```
#Frequency of trips
table(js_data1a$fq5trip)
#Removing na's
js_data1a <- subset(js_data1a, js_data1a$fq5trip != 9)
#Removing missing values
js_data1a <- subset(js_data1a, js_data1a$fq5trip != 8)
#Combining categories "several times a month" and "at least once in
    month" to one category
js_data1a$fq5trip[js_data1a$fq5trip == 1 ] <- 2
#Change order for better undrstanding
js_data1a$fq5trip[js_data1a$fq5trip == 5] <- 0
js_data1a$fq5trip[js_data1a$fq5trip == 4] <- 1
js_data1a$fq5trip[js_data1a$fq5trip == 2] <- 4
js_data1a$fq5trip[js_data1a$fq5trip == 3] <- 2
js_data1a$fq5trip[js_data1a$fq5trip == 4] <- 3
table(js_data1a$fq5trip)

#Rating of health condition
table(js_data1a$op5hlthz)
#Removing nas
js_data1a <- subset(js_data1a, js_data1a$op5hlthz != 9)
#Change order for better undrstanding
js_data1a$op5hlthz[js_data1a$op5hlthz == 5] <- 0
js_data1a$op5hlthz[js_data1a$op5hlthz == 1] <- 5
js_data1a$op5hlthz[js_data1a$op5hlthz == 4] <- 1
js_data1a$op5hlthz[js_data1a$op5hlthz == 2] <- 4
js_data1a$op5hlthz[js_data1a$op5hlthz == 3] <- 2
js_data1a$op5hlthz[js_data1a$op5hlthz == 4] <- 3
js_data1a$op5hlthz[js_data1a$op5hlthz == 5] <- 4
table(js_data1a$op5hlthz)

#Happiness (answer to question: Are you happy)
table(js_data1a$op5happz)
#Removing na's
js_data1a <- subset(js_data1a, js_data1a$op5happz != 9)
#Combining last (poorest) categories
js_data1a$op5happz[js_data1a$op5happz == 5] <- 4
#Change order for better undrstanding
js_data1a$op5happz[js_data1a$op5happz == 4] <- 0
js_data1a$op5happz[js_data1a$op5happz == 1] <- 4
js_data1a$op5happz[js_data1a$op5happz == 3] <- 1
js_data1a$op5happz[js_data1a$op5happz == 4] <- 3
table(js_data1a$op5happz)

#Progress of financial situation in last years
table(js_data1a$op3ecn3a)
#Change order for better undrstanding
js_data1a$op3ecn3a[js_data1a$op3ecn3a == 2] <- 0
js_data1a$op3ecn3a[js_data1a$op3ecn3a == 1] <- 2
js_data1a$op3ecn3a[js_data1a$op3ecn3a == 3] <- 1

#Believe in possiblities to improve standard of living
table(js_data1a$op5chnca)
```

```
#Removing na's
js_data1a <- subset(js_data1a, js_data1a$op5chnca != 9)
#Change order for better undrstanding
js_data1a$op5chnca[js_data1a$op5chnca == 5] <- 0
js_data1a$op5chnca[js_data1a$op5chnca == 1] <- 5
js_data1a$op5chnca[js_data1a$op5chnca == 4] <- 1
js_data1a$op5chnca[js_data1a$op5chnca == 2] <- 4
js_data1a$op5chnca[js_data1a$op5chnca == 3] <- 2
js_data1a$op5chnca[js_data1a$op5chnca == 4] <- 3
js_data1a$op5chnca[js_data1a$op5chnca == 5] <- 4
table(js_data1a$op5chnca)

#Self-reported level in society
table(js_data1a$op5levk)
#Removing na's
js_data1a <- subset(js_data1a, js_data1a$op5levk != 9)
#Change order for better undrstanding
js_data1a$op5levk[js_data1a$op5levk == 5] <- 0
js_data1a$op5levk[js_data1a$op5levk == 1] <- 5
js_data1a$op5levk[js_data1a$op5levk == 4] <- 1
js_data1a$op5levk[js_data1a$op5levk == 2] <- 4
js_data1a$op5levk[js_data1a$op5levk == 3] <- 2
js_data1a$op5levk[js_data1a$op5levk == 4] <- 3
js_data1a$op5levk[js_data1a$op5levk == 5] <- 4
table(js_data1a$op5levk)

#Anxiety about one's economic situation in the future
table(js_data1a$axecnsf)
#Removing na's
js_data1a <- subset(js_data1a, js_data1a$axecnsf != 9)
#Change order for better undrstanding
js_data1a$axecnsf[js_data1a$axecnsf == 5] <- 0
js_data1a$axecnsf[js_data1a$axecnsf == 1] <- 5
js_data1a$axecnsf[js_data1a$axecnsf == 4] <- 1
js_data1a$axecnsf[js_data1a$axecnsf == 2] <- 4
js_data1a$axecnsf[js_data1a$axecnsf == 3] <- 2
js_data1a$axecnsf[js_data1a$axecnsf == 4] <- 3
js_data1a$axecnsf[js_data1a$axecnsf == 5] <- 4
table(js_data1a$axecnsf)

#Trust in people
table(js_data1a$op3trust)
#Removing na's
js_data1a <- subset(js_data1a, js_data1a$op3trust != 9)
#Change order for better undrstanding
js_data1a$op3trust[js_data1a$op3trust == 2] <- 0
js_data1a$op3trust[js_data1a$op3trust == 1] <- 2
js_data1a$op3trust[js_data1a$op3trust == 3] <- 1
table(js_data1a$op3trust)

#Estimation of OLS model
olsa1 <- lm(st5job ~ ageb + agebSquare + male + never_married +
    divorced + widowed + ccnumttl + graduatedCollege + szincomx +
```

```
    xjobhwkLog + executive + temporary + dispatched + self_employed +
     family_worker + GovermentAgency + szcmttl + st5hlthy + st5areay
    + st5leisy + civic + memhobby + fq7frsee + fqsport + fq5read +
    hrtv + fq5trip + op5hlthz  + op5happz + op3ecn3a + op5chnca +
    op5levk + axecnsf + op3trust ,  data=js_data1a)
summary(olsa1)

#Breusch−Pegan test
bptest(olsa1)
#Result: 53.0715, df = 34, p−value = 0.01967
#Estimation of model with heteroskedasticity robust standard errors
coeftest(olsa1, df = Inf, vcov = vcovHC(olsa1, type="HC"))


#Models without irrelvant variables
#Frequency of watching tv least signfivant
olsa1a <− lm(st5job ~ ageb + agebSquare + male + never_married +
    divorced + widowed + ccnumttl + graduatedCollege + szincomx +
    xjobhwkLog + executive + temporary + dispatched + self_employed +
     family_worker + GovermentAgency + szcmttl + st5hlthy + st5areay
    + st5leisy + civic + memhobby + fq7frsee + fqsport + fq5read +
    fq5trip + op5hlthz  + op5happz + op3ecn3a + op5chnca + op5levk +
    axecnsf + op3trust ,  data=js_data1a)
summary(olsa1a)
#Adjusted Rquare higher
coeftest(olsa1a, df = Inf, vcov = vcovHC(olsa1a, type="HC"))
#Dummy variable temporary now least significant
olsa1b <− lm(st5job ~ ageb + agebSquare + male + never_married +
    divorced + widowed + ccnumttl + graduatedCollege + szincomx +
    xjobhwkLog + executive + dispatched + self_employed + family_
    worker + GovermentAgency + szcmttl + st5hlthy + st5areay +
    st5leisy + civic + memhobby + fq7frsee + fqsport + fq5read +
    fq5trip + op5hlthz  + op5happz + op3ecn3a + op5chnca + op5levk +
    axecnsf + op3trust ,  data=js_data1a)
summary(olsa1b)
#Adjusted Rquare higher
coeftest(olsa1b, df = Inf, vcov = vcovHC(olsa1b, type="HC"))
#Number of children now least significant
olsa1c <− lm(st5job ~ ageb + agebSquare + male + never_married +
    divorced + widowed + graduatedCollege + szincomx + xjobhwkLog +
    executive + dispatched + self_employed + family_worker +
    GovermentAgency + szcmttl + st5hlthy + st5areay + st5leisy +
    civic + memhobby + fq7frsee + fqsport + fq5read + fq5trip +
    op5hlthz  + op5happz + op3ecn3a + op5chnca + op5levk + axecnsf +
    op3trust ,  data=js_data1a)
summary(olsa1c)
#Adjusted Rquare higher
coeftest(olsa1c, df = Inf, vcov = vcovHC(olsa1c, type="HC"))
#Frequency of doing sport now least significant
olsa1d <− lm(st5job ~ ageb + agebSquare + male + never_married +
    divorced + widowed + graduatedCollege + szincomx + xjobhwkLog +
    executive + dispatched + self_employed + family_worker +
    GovermentAgency + szcmttl + st5hlthy + st5areay + st5leisy +
```

```
    civic + memhobby + fq7frsee  + fq5read + fq5trip + op5hlthz  +
    op5happz + op3ecn3a + op5chnca + op5levk + axecnsf + op3trust,
    data=js_data1a)
summary(olsa1d)
#Adjusted Rquare higher
coeftest(olsa1d, df = Inf, vcov = vcovHC(olsa1d, type="HC"))
#Dummy varaible dispatsched now least signficant
olsa1e <- lm(st5job ~ ageb + agebSquare + male + never_married +
    divorced + widowed + graduatedCollege + szincomx + xjobhwkLog +
    executive + self_employed + family_worker + GovermentAgency +
    szcmttl + st5hlthy + st5areay + st5leisy + civic + memhobby +
    fq7frsee  + fq5read + fq5trip + op5hlthz  + op5happz + op3ecn3a +
     op5chnca + op5levk + axecnsf + op3trust,  data=js_data1a)
summary(olsa1e)
#Adjusted Rquare higher
coeftest(olsa1e, df = Inf, vcov = vcovHC(olsa1e, type="HC"))
#Dummy variable self-employed now least signifcant
olsa1f <- lm(st5job ~ ageb + agebSquare + male + never_married +
    divorced + widowed + graduatedCollege + szincomx + xjobhwkLog +
    executive  + family_worker + GovermentAgency + szcmttl + st5hlthy
     + st5areay + st5leisy + civic + memhobby + fq7frsee  + fq5read +
     fq5trip + op5hlthz  + op5happz + op3ecn3a + op5chnca + op5levk +
     axecnsf + op3trust,  data=js_data1a)
summary(olsa1f)
#Adjusted Rquare higher
coeftest(olsa1f, df = Inf, vcov = vcovHC(olsa1f, type="HC"))
#age squae now least significant
olsa1g <- lm(st5job ~ ageb + male + never_married + divorced +
    widowed + graduatedCollege + szincomx + xjobhwkLog + executive  +
     family_worker + GovermentAgency + szcmttl + st5hlthy + st5areay
    + st5leisy + civic + memhobby + fq7frsee  + fq5read + fq5trip +
    op5hlthz  + op5happz + op3ecn3a + op5chnca + op5levk + axecnsf +
    op3trust,  data=js_data1a)
summary(olsa1g)
#Adjusted Rquare higher
coeftest(olsa1g, df = Inf, vcov = vcovHC(olsa1g, type="HC"))
#Satisfaction with health condition now least significant
olsa1h <- lm(st5job ~ ageb + male + never_married + divorced +
    widowed + graduatedCollege + szincomx + xjobhwkLog + executive  +
     family_worker + GovermentAgency + szcmttl + st5areay + st5leisy
    + civic + memhobby + fq7frsee  + fq5read + fq5trip + op5hlthz  +
    op5happz + op3ecn3a + op5chnca + op5levk + axecnsf + op3trust,
    data=js_data1a)
summary(olsa1h)
#Adjusted Rquare higher
coeftest(olsa1h, df = Inf, vcov = vcovHC(olsa1h, type="HC"))
#Commuting time now least significant
olsa1i <- lm(st5job ~ ageb + male + never_married + divorced +
    widowed + graduatedCollege + szincomx + xjobhwkLog + executive  +
     family_worker + GovermentAgency  + st5areay + st5leisy + civic +
     memhobby + fq7frsee  + fq5read + fq5trip + op5hlthz  + op5happz
    + op3ecn3a + op5chnca + op5levk + axecnsf + op3trust,  data=js_
    data1a)
```

```
summary(olsa1i)
#Adjusted Rquare higher
coeftest(olsa1i, df = Inf, vcov = vcovHC(olsa1i, type="HC"))
#Membership in hobby organization now least significant
olsa1j <- lm(st5job ~ ageb + male + never_married + divorced +
    widowed + graduatedCollege + szincomx + xjobhwkLog + executive  +
     family_worker + GovermentAgency  + st5areay + st5leisy + civic +
     fq7frsee  + fq5read + fq5trip + op5hlthz  + op5happz + op3ecn3a
    + op5chnca + op5levk + axecnsf + op3trust,  data=js_data1a)
summary(olsa1j)
#Adjusted Rquare higher
coeftest(olsa1j, df = Inf, vcov = vcovHC(olsa1j, type="HC"))
#Working in governmental agency now least signifcant
olsa1k <- lm(st5job ~ ageb + male + never_married + divorced +
    widowed + graduatedCollege + szincomx + xjobhwkLog + executive  +
     family_worker  + st5areay + st5leisy + civic + fq7frsee  +
    fq5read + fq5trip + op5hlthz  + op5happz + op3ecn3a + op5chnca +
    op5levk + axecnsf + op3trust,  data=js_data1a)
summary(olsa1k)
#Adjusted Rquare same
coeftest(olsa1k, df = Inf, vcov = vcovHC(olsa1k, type="HC"))
#College graduation now least signifcant
olsa1l <- lm(st5job ~ ageb + male + never_married + divorced +
    widowed  + szincomx + xjobhwkLog + executive  + family_worker  +
    st5areay + st5leisy + civic + fq7frsee  + fq5read + fq5trip +
    op5hlthz  + op5happz + op3ecn3a + op5chnca + op5levk + axecnsf +
    op3trust,  data=js_data1a)
summary(olsa1l)
#Adjusted Rquare higher
coeftest(olsa1l, df = Inf, vcov = vcovHC(olsa1l, type="HC"))
#Age now least significant
olsa1m <- lm(st5job ~  male + never_married + divorced + widowed  +
    szincomx + xjobhwkLog + executive  + family_worker  + st5areay +
    st5leisy + civic + fq7frsee  + fq5read + fq5trip + op5hlthz  +
    op5happz + op3ecn3a + op5chnca + op5levk + axecnsf + op3trust,
    data=js_data1a)
summary(olsa1m)
#Adjusted Rquare higher
coeftest(olsa1m, df = Inf, vcov = vcovHC(olsa1m, type="HC"))
#Being divorced now least significant
olsa1n <- lm(st5job ~  male + never_married + widowed  + szincomx +
    xjobhwkLog + executive  + family_worker  + st5areay + st5leisy +
    civic + fq7frsee  + fq5read + fq5trip + op5hlthz  + op5happz +
    op3ecn3a + op5chnca + op5levk + axecnsf + op3trust,  data=js_
    data1a)
summary(olsa1n)
#Stopping at previous model (olsa1m) since adjusted r square for
    olsa1n is lower
summary(olsa1m)

vif(olsa1m)

#Descriptive Statistics for general models subsample A
```

```
summary(js_data1a$st5job)
se(js_data1a$st5job)
summary(js_data1a$ageb)
se(js_data1a$ageb)
summary(js_data1a$male)
se(js_data1a$male)
summary(js_data1a$never_married)
se(js_data1a$never_married)
summary(js_data1a$divorced)
se(js_data1a$divorced)
summary(js_data1a$widowed)
se(js_data1a$widowed)
summary(js_data1a$ccnumttl)
se(js_data1a$ccnumttl)
summary(js_data1a$graduatedCollege)
se(js_data1a$graduatedCollege)
summary(js_data1a$szincomx)
se(js_data1a$szincomx)
summary(js_data1a$xjobhwk)
se(js_data1a$xjobhwk)
summary(js_data1a$executive)
se(js_data1a$executive)
summary(js_data1a$temporary)
se(js_data1a$temporary)
summary(js_data1a$dispatched)
se(js_data1a$dispatched)
summary(js_data1a$self_employed)
se(js_data1a$self_employed)
summary(js_data1a$family_worker)
se(js_data1a$family_worker)
summary(js_data1a$GovermentAgency)
se(js_data1a$GovermentAgency)
summary(js_data1a$szcmttl)
se(js_data1a$szcmttl)
summary(js_data1a$st5hlthy)
se(js_data1a$st5hlthy)
summary(js_data1a$st5areay)
se(js_data1a$st5areay)
summary(js_data1a$st5leisy)
se(js_data1a$st5leisy)
summary(js_data1a$civic)
se(js_data1a$civic)
summary(js_data1a$memhobby)
se(js_data1a$memhobby)
summary(js_data1a$fq7frsee)
se(js_data1a$fq7frsee)
summary(js_data1a$fqsport)
se(js_data1a$fqsport)
summary(js_data1a$fq5read)
se(js_data1a$fq5read)
summary(js_data1a$hrtv)
se(js_data1a$hrtv)
summary(js_data1a$fq5trip)
```

```
se(js_data1a$fq5trip)
summary(js_data1a$fq5trip)
se(js_data1a$fq5trip)
summary(js_data1a$op5hlthz)
se(js_data1a$op5hlthz)
summary(js_data1a$op5happz)
se(js_data1a$op5happz)
summary(js_data1a$op3ecn3a)
se(js_data1a$op3ecn3a)
summary(js_data1a$op5chnca)
se(js_data1a$op5chnca)
summary(js_data1a$op5levk)
se(js_data1a$op5levk)
summary(js_data1a$axecnsf)
se(js_data1a$axecnsf)
summary(js_data1a$op3trust)
se(js_data1a$op3trust)


#Estimation of regression for famale sample
js_data1aFem <- js_data1a
#Ereasing male sample
js_data1aFem <- subset(js_data1aFem, js_data1aFem$sexa != 1)
#Estimation OLS
olsa1aFem <- lm(st5job ~ ageb + agebSquare  + never_married +
    divorced + widowed + ccnumttl + graduatedCollege + szincomx +
    xjobhwkLog + executive + temporary + dispatched + self_employed +
     family_worker + GovermentAgency + szcmttl + st5hlthy + st5areay
    + st5leisy + civic + memhobby + fq7frsee + fqsport + fq5read +
    fq5trip + op5hlthz  + op5happz + op3ecn3a + op5chnca + op5levk +
    axecnsf + op3trust, js_data1aFem)
summary(olsa1aFem)
coeftest(olsa1aFem, df = Inf, vcov = vcovHC(olsa1Fem, type="HC"))


#Models without irrelvant variables
#being divorced least signfivant
olsa1aFema <- lm(st5job ~ ageb + agebSquare  + never_married  +
    widowed + ccnumttl + graduatedCollege + szincomx + xjobhwkLog +
    executive + temporary + dispatched + self_employed + family_
    worker + GovermentAgency + szcmttl + st5hlthy + st5areay +
    st5leisy + civic + memhobby + fq7frsee + fqsport + fq5read +
    fq5trip + op5hlthz  + op5happz + op3ecn3a + op5chnca + op5levk +
    axecnsf + op3trust, js_data1aFem)
summary(olsa1aFema)
#Adjusted Rquare higher
coeftest(olsa1aFema, df = Inf, vcov = vcovHC(olsa1aFema, type="HC"))
#Doing sports now least significant
olsa1aFemb <- lm(st5job ~ ageb + agebSquare  + never_married  +
    widowed + ccnumttl + graduatedCollege + szincomx + xjobhwkLog +
    executive + temporary + dispatched + self_employed + family_
    worker + GovermentAgency + szcmttl + st5hlthy + st5areay +
    st5leisy + civic + memhobby + fq7frsee + fq5read + fq5trip +
```

```
    op5hlthz  + op5happz + op3ecn3a + op5chnca + op5levk + axecnsf +
    op3trust , js_data1aFem)
summary(olsa1aFemb)
#Adjusted Rquare higher
coeftest(olsa1aFemb, df = Inf, vcov = vcovHC(olsa1aFemb, type="HC"))
#Commuting time no least significant
olsa1aFemc <- lm(st5job ~ ageb + agebSquare  + never_married  +
    widowed + ccnumttl + graduatedCollege + szincomx + xjobhwkLog +
    executive + temporary + dispatched + self_employed + family_
    worker + GovermentAgency  + st5hlthy + st5areay + st5leisy +
    civic + memhobby + fq7frsee + fq5read + fq5trip + op5hlthz  +
    op5happz + op3ecn3a + op5chnca + op5levk + axecnsf + op3trust , js
    _data1aFem)
summary(olsa1aFemc)
#Adjusted Rquare higher
coeftest(olsa1aFemc, df = Inf, vcov = vcovHC(olsa1aFemc, type="HC"))
#Dummy variable self-employed now least significant
olsa1aFemd <- lm(st5job ~ ageb + agebSquare  + never_married  +
    widowed + ccnumttl + graduatedCollege + szincomx + xjobhwkLog +
    executive + temporary + dispatched  + family_worker +
    GovermentAgency  + st5hlthy + st5areay + st5leisy + civic +
    memhobby + fq7frsee + fq5read + fq5trip + op5hlthz  + op5happz +
    op3ecn3a + op5chnca + op5levk + axecnsf + op3trust , js_data1aFem)
summary(olsa1aFemd)
#Adjusted Rquare higher
coeftest(olsa1aFemd, df = Inf, vcov = vcovHC(olsa1aFemd, type="HC"))
#Being a dispatched worker now least signifivant
olsa1aFeme <- lm(st5job ~ ageb + agebSquare  + never_married  +
    widowed + ccnumttl + graduatedCollege + szincomx + xjobhwkLog +
    executive + temporary + family_worker + GovermentAgency  +
    st5hlthy + st5areay + st5leisy + civic + memhobby + fq7frsee +
    fq5read + fq5trip + op5hlthz  + op5happz + op3ecn3a + op5chnca +
    op5levk + axecnsf + op3trust , js_data1aFem)
summary(olsa1aFeme)
#Adjusted Rquare higher
coeftest(olsa1aFeme, df = Inf, vcov = vcovHC(olsa1aFeme, type="HC"))
#Being widowed now least significant
olsa1aFemf <- lm(st5job ~ ageb + agebSquare  + never_married +
    ccnumttl + graduatedCollege + szincomx + xjobhwkLog + executive +
     temporary + family_worker + GovermentAgency  + st5hlthy +
    st5areay + st5leisy + civic + memhobby + fq7frsee + fq5read +
    fq5trip + op5hlthz  + op5happz + op3ecn3a + op5chnca + op5levk +
    axecnsf + op3trust , js_data1aFem)
summary(olsa1aFemf)
#Adjusted Rquare higher
coeftest(olsa1aFemf, df = Inf, vcov = vcovHC(olsa1aFemf, type="HC"))
#Satisfaction with leisure now least significnat
olsa1aFemg <- lm(st5job ~ ageb + agebSquare  + never_married +
    ccnumttl + graduatedCollege + szincomx + xjobhwkLog + executive +
     temporary + family_worker + GovermentAgency  + st5hlthy +
    st5areay + civic + memhobby + fq7frsee + fq5read + fq5trip +
    op5hlthz  + op5happz + op3ecn3a + op5chnca + op5levk + axecnsf +
    op3trust , js_data1aFem)
```

```r
summary(olsa1aFemg)
#Adjusted Rquare higher
coeftest(olsa1aFemg, df = Inf, vcov = vcovHC(olsa1aFemg, type="HC"))
#Self-repeorted healt condition now least significant
olsa1aFemh <- lm(st5job ~ ageb + agebSquare  + never_married +
    ccnumttl + graduatedCollege + szincomx + xjobhwkLog + executive +
     temporary + family_worker + GovermentAgency  + st5hlthy +
    st5areay + civic + memhobby + fq7frsee + fq5read + fq5trip  +
    op5happz + op3ecn3a + op5chnca + op5levk + axecnsf + op3trust, js
    _data1aFem)
summary(olsa1aFemh)
#Adjusted Rquare higher
coeftest(olsa1aFemh, df = Inf, vcov = vcovHC(olsa1aFemh, type="HC"))
#Number of books now least significant
olsa1aFemi <- lm(st5job ~ ageb + agebSquare  + never_married +
    ccnumttl + graduatedCollege + szincomx + xjobhwkLog + executive +
     temporary + family_worker + GovermentAgency  + st5hlthy +
    st5areay + civic + memhobby + fq7frsee  + fq5trip  + op5happz +
    op3ecn3a + op5chnca + op5levk + axecnsf + op3trust, js_data1aFem)
summary(olsa1aFemi)
#Adjusted Rquare higher
coeftest(olsa1aFemi, df = Inf, vcov = vcovHC(olsa1aFemi, type="HC"))
#Age square now least significant
olsa1aFemj <- lm(st5job ~ ageb  + never_married + ccnumttl +
    graduatedCollege + szincomx + xjobhwkLog + executive + temporary
    + family_worker + GovermentAgency  + st5hlthy + st5areay + civic
    + memhobby + fq7frsee  + fq5trip  + op5happz + op3ecn3a +
    op5chnca + op5levk + axecnsf + op3trust, js_data1aFem)
summary(olsa1aFemj)
#Adjusted Rquare higher
coeftest(olsa1aFemj, df = Inf, vcov = vcovHC(olsa1aFemj, type="HC"))
#Seeing friends now least significant
olsa1aFemk <- lm(st5job ~ ageb  + never_married + ccnumttl +
    graduatedCollege + szincomx + xjobhwkLog + executive + temporary
    + family_worker + GovermentAgency  + st5hlthy + st5areay + civic
    + memhobby + fq5trip  + op5happz + op3ecn3a + op5chnca + op5levk
    + axecnsf + op3trust, js_data1aFem)
summary(olsa1aFemk)
#Adjusted Rquare higher
coeftest(olsa1aFemk, df = Inf, vcov = vcovHC(olsa1aFemk, type="HC"))
#Self-reported social status now least significant
olsa1aFeml <- lm(st5job ~ ageb  + never_married + ccnumttl +
    graduatedCollege + szincomx + xjobhwkLog + executive + temporary
    + family_worker + GovermentAgency  + st5hlthy + st5areay + civic
    + memhobby + fq5trip  + op5happz + op3ecn3a + op5chnca  + axecnsf
     + op3trust, js_data1aFem)
summary(olsa1aFeml)
#Adjusted Rquare higher
coeftest(olsa1aFeml, df = Inf, vcov = vcovHC(olsa1aFeml, type="HC"))
#Satisfaction with health now least significant
olsa1aFemm <- lm(st5job ~ ageb  + never_married + ccnumttl +
    graduatedCollege + szincomx + xjobhwkLog + executive + temporary
    + family_worker + GovermentAgency  + st5areay + civic + memhobby
```

```
    + fq5trip  + op5happz + op3ecn3a + op5chnca  + axecnsf + op3trust
    , js_data1aFem)
summary(olsa1aFemm)
#Adjusted Rquare higher
coeftest(olsa1aFemm, df = Inf, vcov = vcovHC(olsa1aFemm, type="HC"))
#Age now least significant
olsa1aFemn <- lm(st5job ~ never_married + ccnumttl + graduatedCollege
     + szincomx + xjobhwkLog + executive + temporary + family_worker
    + GovermentAgency  + st5areay + civic + memhobby + fq5trip  +
    op5happz + op3ecn3a + op5chnca  + axecnsf + op3trust, js_
    data1aFem)
summary(olsa1aFemn)
#Adjusted Rquare higher
coeftest(olsa1aFemn, df = Inf, vcov = vcovHC(olsa1aFemn, type="HC"))
#Frequency of trips now least significant
olsa1aFemo <- lm(st5job ~ never_married + ccnumttl + graduatedCollege
     + szincomx + xjobhwkLog + executive + temporary + family_worker
    + GovermentAgency  + st5areay + civic + memhobby  + op5happz +
    op3ecn3a + op5chnca  + axecnsf + op3trust, js_data1aFem)
summary(olsa1aFemo)
#Adjusted Rquare higher
coeftest(olsa1aFemo, df = Inf, vcov = vcovHC(olsa1aFemo, type="HC"))
#Graduation from college now least significant
olsa1aFemp <- lm(st5job ~ never_married + ccnumttl  + szincomx +
    xjobhwkLog + executive + temporary + family_worker +
    GovermentAgency  + st5areay + civic + memhobby  + op5happz +
    op3ecn3a + op5chnca  + axecnsf + op3trust, js_data1aFem)
summary(olsa1aFemp)
#Adjusted Rquare higher
coeftest(olsa1aFemp, df = Inf, vcov = vcovHC(olsa1aFemp, type="HC"))
#Membership in hobby organization now least significant
olsa1aFemq <- lm(st5job ~ never_married + ccnumttl  + szincomx +
    xjobhwkLog + executive + temporary + family_worker +
    GovermentAgency  + st5areay + civic  + op5happz + op3ecn3a +
    op5chnca  + axecnsf + op3trust, js_data1aFem)
summary(olsa1aFemq)
#Adjusted Rquare higher
coeftest(olsa1aFemq, df = Inf, vcov = vcovHC(olsa1aFemq, type="HC"))
#Number of children now least significnat
olsa1aFemr <- lm(st5job ~ never_married  + szincomx + xjobhwkLog +
    executive + temporary + family_worker + GovermentAgency  +
    st5areay + civic  + op5happz + op3ecn3a + op5chnca  + axecnsf +
    op3trust, js_data1aFem)
summary(olsa1aFemr)
#Stopping at previous model (olsa1aFemq) since adjusted r square for
    olsa1aFemr is lower
#Dummy variable temporary second least significant

#Descriptive Statistics for female model of subsample A
summary(js_data1aFem$st5job)
se(js_data1aFem$st5job)
summary(js_data1aFem$ageb)
se(js_data1aFem$ageb)
```

```
summary(js_data1aFem$male)
se(js_data1aFem$male)
summary(js_data1aFem$never_married)
se(js_data1aFem$never_married)
summary(js_data1aFem$divorced)
se(js_data1aFem$divorced)
summary(js_data1aFem$widowed)
se(js_data1aFem$widowed)
summary(js_data1aFem$ccnumttl)
se(js_data1aFem$ccnumttl)
summary(js_data1aFem$graduatedCollege)
se(js_data1aFem$graduatedCollege)
summary(js_data1aFem$szincomx)
se(js_data1aFem$szincomx)
summary(js_data1aFem$xjobhwk)
se(js_data1aFem$xjobhwk)
summary(js_data1aFem$executive)
se(js_data1aFem$executive)
summary(js_data1aFem$temporary)
se(js_data1aFem$temporary)
summary(js_data1aFem$dispatched)
se(js_data1aFem$dispatched)
summary(js_data1aFem$self_employed)
se(js_data1aFem$self_employed)
summary(js_data1aFem$family_worker)
se(js_data1aFem$family_worker)
summary(js_data1aFem$GovermentAgency)
se(js_data1aFem$GovermentAgency)
summary(js_data1aFem$szcmttl)
se(js_data1aFem$szcmttl)
summary(js_data1aFem$st5hlthy)
se(js_data1aFem$st5hlthy)
summary(js_data1aFem$st5areay)
se(js_data1aFem$st5areay)
summary(js_data1aFem$st5leisy)
se(js_data1aFem$st5leisy)
summary(js_data1aFem$civic)
se(js_data1aFem$civic)
summary(js_data1aFem$memhobby)
se(js_data1aFem$memhobby)
summary(js_data1aFem$fq7frsee)
se(js_data1aFem$fq7frsee)
summary(js_data1aFem$fqsport)
se(js_data1aFem$fqsport)
summary(js_data1aFem$fq5read)
se(js_data1aFem$fq5read)
summary(js_data1aFem$hrtv)
se(js_data1aFem$hrtv)
summary(js_data1aFem$fq5trip)
se(js_data1aFem$fq5trip)
summary(js_data1aFem$op5hlthz)
se(js_data1aFem$op5hlthz)
summary(js_data1aFem$op5happz)
```

```
se(js_data1aFem$op5happz)
summary(js_data1aFem$op3ecn3a)
se(js_data1aFem$op3ecn3a)
summary(js_data1aFem$op5chnca)
se(js_data1aFem$op5chnca)
summary(js_data1aFem$op5levk)
se(js_data1aFem$op5levk)
summary(js_data1aFem$axecnsf)
se(js_data1aFem$axecnsf)
summary(js_data1aFem$op3trust)
se(js_data1aFem$op3trust)


#Estimation of regression for male sample
js_data1aMal <- js_data1a
table(js_data1aMal$sexa)
#Ereasing male sample
js_data1aMal <- subset(js_data1aMal,js_data1aMal$sexa != 0)
#Estimation OLS
olsa1Mal <- lm(st5job ~ ageb + agebSquare  + never_married + divorced + widowed + ccnumttl + graduatedCollege + szincomx + xjobhwkLog + executive + temporary + dispatched + self_employed + family_worker + GovermentAgency + szcmttl + st5hlthy + st5areay + st5leisy + civic + memhobby + fq7frsee + fqsport + fq5read + fq5trip + op5hlthz  + op5happz + op3ecn3a + op5chnca + op5levk + axecnsf + op3trust ,  js_data1aMal)
summary(olsa1Mal)
coeftest(olsa1Mal, df = Inf, vcov = vcovHC(olsa1Mal, type="HC"))
#Working for Government agenc least significant
olsa1Mala <- lm(st5job ~ ageb + agebSquare  + never_married + divorced + widowed + ccnumttl + graduatedCollege + szincomx + xjobhwkLog + executive + temporary + dispatched + self_employed + family_worker  + szcmttl + st5hlthy + st5areay + st5leisy + civic + memhobby + fq7frsee + fqsport + fq5read + fq5trip + op5hlthz  + op5happz + op3ecn3a + op5chnca + op5levk + axecnsf + op3trust ,  js_data1aMal)
summary(olsa1Mala)
#Adjusted Rquare higher
coeftest(olsa1Mala, df = Inf, vcov = vcovHC(olsa1Mala, type="HC"))
#Being family worker now least significant
olsa1Malb <- lm(st5job ~ ageb + agebSquare  + never_married + divorced + widowed + ccnumttl + graduatedCollege + szincomx + xjobhwkLog + executive + temporary + dispatched + self_employed + szcmttl + st5hlthy + st5areay + st5leisy + civic + memhobby + fq7frsee + fqsport + fq5read + fq5trip + op5hlthz  + op5happz + op3ecn3a + op5chnca + op5levk + axecnsf + op3trust ,  js_data1aMal)
summary(olsa1Malb)
#Adjusted Rquare higher
coeftest(olsa1Malb, df = Inf, vcov = vcovHC(olsa1Malb, type="HC"))
#Graduation from college now least significant
olsa1Malc <- lm(st5job ~ ageb + agebSquare  + never_married + divorced + widowed + ccnumttl  + szincomx + xjobhwkLog +
```

```
    executive + temporary + dispatched + self_employed  + szcmttl +
    st5hlthy + st5areay + st5leisy + civic + memhobby + fq7frsee +
    fqsport + fq5read + fq5trip + op5hlthz  + op5happz + op3ecn3a +
    op5chnca + op5levk + axecnsf + op3trust ,  js_data1aMal)
summary(olsa1Malc)
#Adjusted Rquare higher
coeftest(olsa1Malc, df = Inf, vcov = vcovHC(olsa1Malc, type="HC"))
#being temorary employed now least significant
olsa1Mald <- lm(st5job ~ ageb + agebSquare  + never_married +
    divorced + widowed + ccnumttl  + szincomx + xjobhwkLog +
    executive  + dispatched + self_employed  + szcmttl + st5hlthy +
    st5areay + st5leisy + civic + memhobby + fq7frsee + fqsport +
    fq5read + fq5trip + op5hlthz  + op5happz + op3ecn3a + op5chnca +
    op5levk + axecnsf + op3trust,  js_data1aMal)
summary(olsa1Mald)
#Adjusted Rquare higher
coeftest(olsa1Mald, df = Inf, vcov = vcovHC(olsa1Mald, type="HC"))
#Doing sports now least significant
olsa1Male <- lm(st5job ~ ageb + agebSquare  + never_married +
    divorced + widowed + ccnumttl  + szincomx + xjobhwkLog +
    executive  + dispatched + self_employed  + szcmttl + st5hlthy +
    st5areay + st5leisy + civic + memhobby + fq7frsee  + fq5read +
    fq5trip + op5hlthz  + op5happz + op3ecn3a + op5chnca + op5levk +
    axecnsf + op3trust,  js_data1aMal)
summary(olsa1Male)
#Adjusted Rquare higher
coeftest(olsa1Male, df = Inf, vcov = vcovHC(olsa1Male, type="HC"))
#Believeing in opportunities to change one's economic situation now
    least significant
olsa1Malf <- lm(st5job ~ ageb + agebSquare  + never_married +
    divorced + widowed + ccnumttl  + szincomx + xjobhwkLog +
    executive  + dispatched + self_employed  + szcmttl + st5hlthy +
    st5areay + st5leisy + civic + memhobby + fq7frsee  + fq5read +
    fq5trip + op5hlthz  + op5happz + op3ecn3a  + op5levk + axecnsf +
    op3trust,  js_data1aMal)
summary(olsa1Malf)
#Adjusted Rquare higher
coeftest(olsa1Malf, df = Inf, vcov = vcovHC(olsa1Malf, type="HC"))
#Commuting time now least significant
olsa1Malg <- lm(st5job ~ ageb + agebSquare  + never_married +
    divorced + widowed + ccnumttl  + szincomx + xjobhwkLog +
    executive  + dispatched + self_employed   + st5hlthy + st5areay +
     st5leisy + civic + memhobby + fq7frsee  + fq5read + fq5trip +
    op5hlthz  + op5happz + op3ecn3a  + op5levk + axecnsf + op3trust,
     js_data1aMal)
summary(olsa1Malg)
#Adjusted Rquare higher
coeftest(olsa1Malg, df = Inf, vcov = vcovHC(olsa1Malg, type="HC"))
#Number of children now least significant
olsa1Malh <- lm(st5job ~ ageb + agebSquare  + never_married +
    divorced + widowed + szincomx + xjobhwkLog + executive  +
    dispatched + self_employed   + st5hlthy + st5areay + st5leisy +
    civic + memhobby + fq7frsee  + fq5read + fq5trip + op5hlthz  +
```

```
    op5happz + op3ecn3a  + op5levk + axecnsf + op3trust ,  js_
    data1aMal)
summary(olsa1Malh)
#Adjusted Rquare higher
coeftest(olsa1Malh , df = Inf , vcov = vcovHC(olsa1Malh , type="HC"))
#Satisfaction with ones health now least significant
olsa1Mali <- lm(st5job ~ ageb + agebSquare  + never_married +
    divorced + widowed + szincomx + xjobhwkLog + executive  +
    dispatched + self_employed + st5areay + st5leisy + civic +
    memhobby + fq7frsee  + fq5read + fq5trip + op5hlthz  + op5happz +
     op3ecn3a  + op5levk + axecnsf + op3trust ,  js_data1aMal)
summary(olsa1Mali)
#Adjusted Rquare higher
coeftest(olsa1Mali , df = Inf , vcov = vcovHC(olsa1Mali , type="HC"))
#Being a dispatched worker now least significant
olsa1Malj <- lm(st5job ~ ageb + agebSquare  + never_married +
    divorced + widowed + szincomx + xjobhwkLog + executive + self_
    employed + st5areay + st5leisy + civic + memhobby + fq7frsee  +
    fq5read + fq5trip + op5hlthz  + op5happz + op3ecn3a  + op5levk +
    axecnsf + op3trust ,  js_data1aMal)
summary(olsa1Malj)
#Adjusted Rquare higher
coeftest(olsa1Malj , df = Inf , vcov = vcovHC(olsa1Malj , type="HC"))
#Age square now most significant
olsa1Malk <- lm(st5job ~ ageb + never_married + divorced + widowed +
    szincomx + xjobhwkLog + executive + self_employed + st5areay +
    st5leisy + civic + memhobby + fq7frsee  + fq5read + fq5trip +
    op5hlthz  + op5happz + op3ecn3a  + op5levk + axecnsf + op3trust ,
     js_data1aMal)
summary(olsa1Malk)
#Adjusted Rquare higher
coeftest(olsa1Malk , df = Inf , vcov = vcovHC(olsa1Malk , type="HC"))
#Membership in hobby organization now least significant
olsa1Mall <- lm(st5job ~ ageb + never_married + divorced + widowed +
    szincomx + xjobhwkLog + executive + self_employed + st5areay +
    st5leisy + civic  + fq7frsee  + fq5read + fq5trip + op5hlthz  +
    op5happz + op3ecn3a  + op5levk + axecnsf + op3trust ,  js_
    data1aMal)
summary(olsa1Mall)
#Adjusted Rquare higher
coeftest(olsa1Mall , df = Inf , vcov = vcovHC(olsa1Mall , type="HC"))
#Satisfaction with one's neigborhood  now least significant
olsa1Malm <- lm(st5job ~ ageb + never_married + divorced + widowed +
    szincomx + xjobhwkLog + executive + self_employed  + st5leisy +
    civic  + fq7frsee  + fq5read + fq5trip + op5hlthz  + op5happz +
    op3ecn3a  + op5levk + axecnsf + op3trust ,  js_data1aMal)
summary(olsa1Malm)
#Adjusted Rquare higher
coeftest(olsa1Malm , df = Inf , vcov = vcovHC(olsa1Malm , type="HC"))
#Being self-employed now least significant
olsa1Maln <- lm(st5job ~ ageb + agebSquare + never_married + divorced
     + widowed  + szincomx + xjobhwkLog + executive + st5leisy +
    civic  + fq7frsee  + fq5read  + fq5trip + op5hlthz  + op5happz  +
```

```
op5chnca + op5levk + axecnsf + op3trust ,  js_data1aMal)
summary(olsa1Maln)
#Adjusted Rquare higher
coeftest(olsa1Maln, df = Inf, vcov = vcovHC(olsa1Maln, type="HC"))
#Believe in opportunity to change one's position  now least significant
olsa1Malo <- lm(st5job ~ ageb + agebSquare + never_married + divorced + widowed  + szincomx + xjobhwkLog + executive + st5leisy + civic  + fq7frsee  + fq5read  + fq5trip + op5hlthz  + op5happz  + op5levk + axecnsf + op3trust ,  js_data1aMal)
summary(olsa1Malo)
#Adjusted Rquare higher
coeftest(olsa1Malo, df = Inf, vcov = vcovHC(olsa1Malo, type="HC"))
#Working hours now least significant
olsa1Malp <- lm(st5job ~ ageb + agebSquare + never_married + divorced + widowed  + szincomx  + executive + st5leisy + civic  + fq7frsee  + fq5read  + fq5trip + op5hlthz  + op5happz  + op5levk + axecnsf + op3trust ,  js_data1aMal)
summary(olsa1Malp)
#Adjusted Rquare higher
coeftest(olsa1Malp, df = Inf, vcov = vcovHC(olsa1Malp, type="HC"))
#Being divorced now least significant
olsa1Malq <- lm(st5job ~ ageb + agebSquare + never_married  + widowed  + szincomx  + executive + st5leisy + civic  + fq7frsee  + fq5read  + fq5trip + op5hlthz  + op5happz  + op5levk + axecnsf + op3trust ,  js_data1aMal)
summary(olsa1Malq)
#Adjusted Rquare higher
coeftest(olsa1Malq, df = Inf, vcov = vcovHC(olsa1Malq, type="HC"))
#Membership in civic organzation now least significant
olsa1Malr <- lm(st5job ~ ageb + agebSquare + never_married  + widowed  + szincomx  + executive + st5leisy  + fq7frsee  + fq5read  + fq5trip + op5hlthz  + op5happz  + op5levk + axecnsf + op3trust ,  js_data1aMal)
summary(olsa1Malr)
#Adjusted Rquare higher
coeftest(olsa1Malr, df = Inf, vcov = vcovHC(olsa1Malr, type="HC"))
#Age and age square now least significant
olsa1Mals <- lm(st5job ~ never_married  + widowed  + szincomx  + executive + st5leisy  + fq7frsee  + fq5read  + fq5trip + op5hlthz  + op5happz  + op5levk + axecnsf + op3trust ,  js_data1aMal)
summary(olsa1Mals)
#Adjusted Rquare higher
coeftest(olsa1Mals, df = Inf, vcov = vcovHC(olsa1Mals, type="HC"))
#Being widowed now least significant
olsa1Malt <- lm(st5job ~  widowed  + szincomx  + executive + st5leisy  + fq7frsee  + fq5read  + fq5trip + op5hlthz  + op5happz  + op5levk + axecnsf + op3trust ,  js_data1aMal)
summary(olsa1Malt)
#Adjusted Rquare higher
coeftest(olsa1Malt, df = Inf, vcov = vcovHC(olsa1Malt, type="HC"))
#Seing friends now least significant
```

```
olsa1Malu <- lm(st5job ~  widowed  + szincomx  + executive + st5leisy
      + fq5read  + fq5trip + op5hlthz  + op5happz  + op5levk +
    axecnsf + op3trust ,  js_data1aMal)
summary(olsa1Malu)
#Stopping at previous model (olsa1aMalt) since adjusted r square for
    olsa1aMalu is lower

#Descriptive Statistics for male model of subsample A
summary(js_data1aMal$st5job)
se(js_data1aMal$st5job)
summary(js_data1aMal$ageb)
se(js_data1aMal$ageb)
summary(js_data1aMal$male)
se(js_data1aMal$male)
summary(js_data1aMal$never_married)
se(js_data1aMal$never_married)
summary(js_data1aMal$divorced)
se(js_data1aMal$divorced)
summary(js_data1aMal$widowed)
se(js_data1aMal$widowed)
summary(js_data1aMal$ccnumttl)
se(js_data1aMal$ccnumttl)
summary(js_data1aMal$graduatedCollege)
se(js_data1aMal$graduatedCollege)
summary(js_data1aMal$szincomx)
se(js_data1aMal$szincomx)
summary(js_data1aMal$xjobhwk)
se(js_data1aMal$xjobhwk)
summary(js_data1aMal$executive)
se(js_data1aMal$executive)
summary(js_data1aMal$temporary)
se(js_data1aMal$temporary)
summary(js_data1aMal$dispatched)
se(js_data1aMal$dispatched)
summary(js_data1aMal$self_employed)
se(js_data1aMal$self_employed)
summary(js_data1aMal$family_worker)
se(js_data1aMal$family_worker)
summary(js_data1aMal$GovermentAgency)
se(js_data1aMal$GovermentAgency)
summary(js_data1aMal$szcmttl)
se(js_data1aMal$szcmttl)
summary(js_data1aMal$st5hlthy)
se(js_data1aMal$st5hlthy)
summary(js_data1aMal$st5areay)
se(js_data1aMal$st5areay)
summary(js_data1aMal$st5leisy)
se(js_data1aMal$st5leisy)
summary(js_data1aMal$civic)
se(js_data1aMal$civic)
summary(js_data1aMal$memhobby)
se(js_data1aMal$memhobby)
summary(js_data1aMal$fq7frsee)
```

```
se(js_data1aMal$fq7frsee)
summary(js_data1aMal$fqsport)
se(js_data1aMal$fqsport)
summary(js_data1aMal$fq5read)
se(js_data1aMal$fq5read)
summary(js_data1aMal$hrtv)
se(js_data1aMal$hrtv)
summary(js_data1aMal$fq5trip)
se(js_data1aMal$fq5trip)
summary(js_data1aMal$op5hlthz)
se(js_data1aMal$op5hlthz)
summary(js_data1aMal$op5happz)
se(js_data1aMal$op5happz)
summary(js_data1aMal$op3ecn3a)
se(js_data1aMal$op3ecn3a)
summary(js_data1aMal$op5chnca)
se(js_data1aMal$op5chnca)
summary(js_data1aMal$op5levk)
se(js_data1aMal$op5levk)
summary(js_data1aMal$axecnsf)
se(js_data1aMal$axecnsf)
summary(js_data1aMal$op3trust)
se(js_data1aMal$op3trust)

stargazer(list(coeftest(olsa1m, df = Inf, vcov = vcovHC(olsa1m, type=
    "HC")), coeftest(olsa1aFemq, df = Inf, vcov = vcovHC(olsa1aFemq,
    type="HC")), coeftest(olsa1Malt, df = Inf, vcov = vcovHC(
    olsa1Malt, type="HC"))))

summary(olsa1m)
summary(olsa1aFemq)
summary(olsa1Malt)

vif(olsa1m)
vif(olsa1aFemq)
vif(olsa1Malt)


#Models derived from Form B
js_data1b <- js_data1

#Self-reported happyness question "in general, are you happy?"
table(js_data1b$op5happe)
#Remove nas
js_data1b <- subset(js_data1b, js_data1b$op5happe != 9 &  js_data1b$
    op5happe !=8)
#Change order for better understanding
js_data1b$op5happe[js_data1b$op5happe == 5] <- 0
js_data1b$op5happe[js_data1b$op5happe == 1] <- 5
js_data1b$op5happe[js_data1b$op5happe == 4] <- 1
js_data1b$op5happe[js_data1b$op5happe == 2] <- 4
js_data1b$op5happe[js_data1b$op5happe == 1] <- 2
js_data1b$op5happe[js_data1b$op5happe == 0] <- 1
```

```
#Progress of financial situation in last years (also for form A)
table(js_data1b$op3ecn3a)
#Change order for better undrstanding
js_data1b$op3ecn3a[js_data1b$op3ecn3a == 2] <- 0
js_data1b$op3ecn3a[js_data1b$op3ecn3a == 1] <- 2
js_data1b$op3ecn3a[js_data1b$op3ecn3a == 3] <- 1

#Believe in possiblities to improve standard of living (also for form A)
table(js_data1b$op5chnca)
#Removing na's
js_data1b <- subset(js_data1b, js_data1b$op5chnca != 9)
#Change order for better undrstanding
js_data1b$op5chnca[js_data1b$op5chnca == 5] <- 0
js_data1b$op5chnca[js_data1b$op5chnca == 1] <- 5
js_data1b$op5chnca[js_data1b$op5chnca == 4] <- 1
js_data1b$op5chnca[js_data1b$op5chnca == 2] <- 4
js_data1b$op5chnca[js_data1b$op5chnca == 3] <- 2
js_data1b$op5chnca[js_data1b$op5chnca == 4] <- 3
js_data1b$op5chnca[js_data1b$op5chnca == 5] <- 4
table(js_data1b$op5chnca)

#Self-reported level in society (also for form A)
table(js_data1b$op5levk)
#Removing na's
js_data1b <- subset(js_data1b, js_data1b$op5levk != 9)
#Change order for better undrstanding
js_data1b$op5levk[js_data1b$op5levk == 5] <- 0
js_data1b$op5levk[js_data1b$op5levk == 1] <- 5
js_data1b$op5levk[js_data1b$op5levk == 4] <- 1
js_data1b$op5levk[js_data1b$op5levk == 2] <- 4
js_data1b$op5levk[js_data1b$op5levk == 3] <- 2
js_data1b$op5levk[js_data1b$op5levk == 4] <- 3
js_data1b$op5levk[js_data1b$op5levk == 5] <- 4
table(js_data1b$op5levk)

#Anxiety about one's economic situation in the future (also for form A)
table(js_data1b$axecnsf)
#Removing na's
js_data1b <- subset(js_data1b, js_data1b$axecnsf != 9)
#Change order for better undrstanding
js_data1b$axecnsf[js_data1b$axecnsf == 5] <- 0
js_data1b$axecnsf[js_data1b$axecnsf == 1] <- 5
js_data1b$axecnsf[js_data1b$axecnsf == 4] <- 1
js_data1b$axecnsf[js_data1b$axecnsf == 2] <- 4
js_data1b$axecnsf[js_data1b$axecnsf == 3] <- 2
js_data1b$axecnsf[js_data1b$axecnsf == 4] <- 3
js_data1b$axecnsf[js_data1b$axecnsf == 5] <- 4
table(js_data1b$axecnsf)

#Soldarity at work place
```

```r
table(js_data1b$opwslpw)
#Removing na's
js_data1b <- subset(js_data1b, js_data1b$opwslpw != 9 & js_data1b$
    opwslpw != 5)
#Changing order for better understanding
js_data1b$opwslpw[js_data1b$opwslpw == 4] <- 0
js_data1b$opwslpw[js_data1b$opwslpw == 1] <- 4
js_data1b$opwslpw[js_data1b$opwslpw == 3] <- 1
js_data1b$opwslpw[js_data1b$opwslpw == 4] <- 3

#Self-reported health condition
table(js_data1b$sfhlcnd)
#Removing nas
js_data1b <- subset(js_data1b, js_data1b$sfhlcnd != 9)
#Change order for better undrstanding
js_data1b$sfhlcnd[js_data1b$sfhlcnd == 5] <- 0
js_data1b$sfhlcnd[js_data1b$sfhlcnd == 1] <- 5
js_data1b$sfhlcnd[js_data1b$sfhlcnd == 4] <- 1
js_data1b$sfhlcnd[js_data1b$sfhlcnd == 2] <- 4
js_data1b$sfhlcnd[js_data1b$sfhlcnd == 3] <- 2
js_data1b$sfhlcnd[js_data1b$sfhlcnd == 4] <- 3
js_data1b$sfhlcnd[js_data1b$sfhlcnd == 5] <- 4
table(js_data1b$sfhlcnd)

#Pain
table(js_data1b$sfintfpn)

#Trust
table(js_data1b$op4trust)
#Removing nas
js_data1b <- subset(js_data1b, js_data1b$op4trust != 9)
#Change order for better undrstanding
js_data1b$op4trust[js_data1b$op4trust == 4] <- 0
js_data1b$op4trust[js_data1b$op4trust == 1] <- 4
js_data1b$op4trust[js_data1b$op4trust == 3] <- 1
js_data1b$op4trust[js_data1b$op4trust == 4] <- 3

#Loneliness
table(js_data1b$lonely)
#Removing nas
js_data1b <- subset(js_data1b, js_data1b$lonely != 9)
#Change order for better understanding
js_data1b$lonely[js_data1b$lonely == 4] <- 0
js_data1b$lonely[js_data1b$lonely == 1] <- 4
js_data1b$lonely[js_data1b$lonely == 3] <- 1
js_data1b$lonely[js_data1b$lonely == 4] <- 3

#Time in nature (woods, ocean, or river)
table(js_data1b$frlxnatr)
js_data1b <- subset(js_data1b, js_data1b$frlxnatr != 9)
#Combining 1,2,3 "to at least once in weak" category
js_data1b$frlxnatr[js_data1b$frlxnatr == 2 | js_data1b$frlxnatr == 1]
    <- 3
```

```
#Changing order for better understanding
js_data1b$frlxnatr[js_data1b$frlxnatr == 6] <- 0
js_data1b$frlxnatr[js_data1b$frlxnatr == 5] <- 1
js_data1b$frlxnatr[js_data1b$frlxnatr == 4] <- 2
table(js_data1b$frlxnatr)

olsb1a <- lm(st5job ~ ageb + agebSquare + male + never_married +
    divorced + widowed + ccnumttl + graduatedCollege + szincomx +
    xjobhwkLog + executive + temporary + dispatched + self_employed +
     family_worker + GovermentAgency + szcmttl + st5hlthy + st5areay
    + st5leisy + civic + memhobby + fq7frsee + fqsport + fq5read +
    frlxnatr + lonely + op4trust + sfintfpn + sfhlcnd + op5happe +
    opwslpw + op3ecn3a + op5chnca + op5levk + axecnsf,  data=js_
    data1b)
summary(olsb1a)
coeftest(olsb1a, df = Inf, vcov = vcovHC(olsb1a, type="HC"))
#Divorced now least significant
olsb1b <- lm(st5job ~ ageb + agebSquare + male + never_married  +
    widowed + ccnumttl + graduatedCollege + szincomx + xjobhwkLog +
    executive + temporary + dispatched + self_employed + family_
    worker + GovermentAgency + szcmttl + st5hlthy + st5areay +
    st5leisy + civic + memhobby + fq7frsee + fqsport + fq5read +
    frlxnatr + lonely + op4trust + sfintfpn + sfhlcnd + op5happe +
    opwslpw + op3ecn3a + op5chnca + op5levk + axecnsf,  data=js_
    data1b)
summary(olsb1b)
#Adjusted Rquare higher
coeftest(olsb1b, df = Inf, vcov = vcovHC(olsb1b, type="HC"))
#Doing sports now least significant
olsb1c <- lm(st5job ~ ageb + agebSquare + male + never_married  +
    widowed + ccnumttl + graduatedCollege + szincomx + xjobhwkLog +
    executive + temporary + dispatched + self_employed + family_
    worker + GovermentAgency + szcmttl + st5hlthy + st5areay +
    st5leisy + civic + memhobby + fq7frsee  + fq5read + frlxnatr +
    lonely + op4trust + sfintfpn + sfhlcnd + op5happe + opwslpw +
    op3ecn3a + op5chnca + op5levk + axecnsf,  data=js_data1b)
summary(olsb1c)
#Adjusted Rquare higher
coeftest(olsb1c, df = Inf, vcov = vcovHC(olsb1c, type="HC"))
#Satisfaction with living area now least significant
olsb1d <- lm(st5job ~ ageb + agebSquare + male + never_married  +
    widowed + ccnumttl + graduatedCollege + szincomx + xjobhwkLog +
    executive + temporary + dispatched + self_employed + family_
    worker + GovermentAgency + szcmttl + st5hlthy  + st5leisy + civic
     + memhobby + fq7frsee  + fq5read + frlxnatr + lonely + op4trust
    + sfintfpn + sfhlcnd + op5happe + opwslpw + op3ecn3a + op5chnca +
     op5levk + axecnsf,  data=js_data1b)
summary(olsb1d)
#Adjusted Rquare higher
coeftest(olsb1d, df = Inf, vcov = vcovHC(olsb1d, type="HC"))
#Being widowed now least signifcant
olsb1e <- lm(st5job ~ ageb + agebSquare + male + never_married  +
    ccnumttl + graduatedCollege + szincomx + xjobhwkLog + executive +
```

```
    temporary + dispatched + self_employed + family_worker +
    GovermentAgency + szcmttl + st5hlthy  + st5leisy + civic +
    memhobby + fq7frsee  + fq5read + frlxnatr + lonely + op4trust +
    sfintfpn + sfhlcnd + op5happe + opwslpw + op3ecn3a + op5chnca +
    op5levk + axecnsf,  data=js_data1b)
summary(olsb1e)
#Adjusted Rquare higher
coeftest(olsb1e, df = Inf, vcov = vcovHC(olsb1e, type="HC"))
#Being self-employed  now least significant
olsb1f <- lm(st5job ~ ageb + agebSquare + male + never_married  +
    ccnumttl + graduatedCollege + szincomx + xjobhwkLog + executive +
    temporary + dispatched  + family_worker + GovermentAgency +
    szcmttl + st5hlthy  + st5leisy + civic + memhobby + fq7frsee  +
    fq5read + frlxnatr + lonely + op4trust + sfintfpn + sfhlcnd +
    op5happe + opwslpw + op3ecn3a + op5chnca + op5levk + axecnsf,
    data=js_data1b)
summary(olsb1f)
#Adjusted Rquare higher
coeftest(olsb1f, df = Inf, vcov = vcovHC(olsb1f, type="HC"))
#Self-reported socia status now least signifcant
olsb1g <- lm(st5job ~ ageb + agebSquare + male + never_married  +
    ccnumttl + graduatedCollege + szincomx + xjobhwkLog + executive +
    temporary + dispatched  + family_worker + GovermentAgency +
    szcmttl + st5hlthy  + st5leisy + civic + memhobby + fq7frsee  +
    fq5read + frlxnatr + lonely + op4trust + sfintfpn + sfhlcnd +
    op5happe + opwslpw + op3ecn3a + op5chnca + axecnsf,  data=js_
    data1b)
summary(olsb1g)
#Adjusted Rquare higher
coeftest(olsb1g, df = Inf, vcov = vcovHC(olsb1g, type="HC"))
#Seeing friends now least signifcant
olsb1h <- lm(st5job ~ ageb + agebSquare + male + never_married  +
    ccnumttl + graduatedCollege + szincomx + xjobhwkLog + executive +
    temporary + dispatched  + family_worker + GovermentAgency +
    szcmttl + st5hlthy  + st5leisy + civic + memhobby + fq5read +
    frlxnatr + lonely + op4trust + sfintfpn + sfhlcnd + op5happe +
    opwslpw + op3ecn3a + op5chnca + axecnsf,  data=js_data1b)
summary(olsb1h)
#Adjusted Rquare higher
coeftest(olsb1h, df = Inf, vcov = vcovHC(olsb1h, type="HC"))
#Self-reported health-condition now least signficant
olsb1i <- lm(st5job ~ ageb + agebSquare + male + never_married  +
    ccnumttl + graduatedCollege + szincomx + xjobhwkLog + executive +
    temporary + dispatched  + family_worker + GovermentAgency +
    szcmttl + st5hlthy  + st5leisy + civic + memhobby + fq5read +
    frlxnatr + lonely + op4trust + sfintfpn+ op5happe + opwslpw +
    op3ecn3a + op5chnca + axecnsf,  data=js_data1b)
summary(olsb1i)
#Adjusted Rquare higher
coeftest(olsb1i, df = Inf, vcov = vcovHC(olsb1i, type="HC"))
#Number of children now least signicant
olsb1j <- lm(st5job ~ ageb + agebSquare + male + never_married  +
    graduatedCollege + szincomx + xjobhwkLog + executive + temporary
```

```
    + dispatched  + family_worker + GovermentAgency + szcmttl +
    st5hlthy  + st5leisy + civic + memhobby + fq5read + frlxnatr +
    lonely + op4trust + sfintfpn+ op5happe + opwslpw + op3ecn3a +
    op5chnca + axecnsf,  data=js_data1b)
summary(olsb1j)
#Adjusted Rquare higher
coeftest(olsb1j, df = Inf, vcov = vcovHC(olsb1j, type="HC"))
#Temporary employemnet now least significant
olsb1k <- lm(st5job ~ ageb + agebSquare + male + never_married  +
    graduatedCollege + szincomx + xjobhwkLog + executive  +
    dispatched  + family_worker + GovermentAgency + szcmttl +
    st5hlthy  + st5leisy + civic + memhobby + fq5read + frlxnatr +
    lonely + op4trust + sfintfpn+ op5happe + opwslpw + op3ecn3a +
    op5chnca + axecnsf,  data=js_data1b)
summary(olsb1k)
#Adjusted Rquare higher
coeftest(olsb1k, df = Inf, vcov = vcovHC(olsb1k, type="HC"))
#Working at Government agency now least significant now least
    significant
olsb1l <- lm(st5job ~ ageb + agebSquare + male + never_married  +
    graduatedCollege + szincomx + xjobhwkLog + executive  +
    dispatched  + family_worker + szcmttl + st5hlthy  + st5leisy +
    civic + memhobby + fq5read + frlxnatr + lonely + op4trust +
    sfintfpn+ op5happe + opwslpw + op3ecn3a + op5chnca + axecnsf,
    data=js_data1b)
summary(olsb1l)
#Adjusted Rquare higher
coeftest(olsb1l, df = Inf, vcov = vcovHC(olsb1l, type="HC"))
#Trust now least significant
olsb1m <- lm(st5job ~ ageb + agebSquare + male + never_married  +
    graduatedCollege + szincomx + xjobhwkLog + executive  +
    dispatched  + family_worker + szcmttl + st5hlthy  + st5leisy +
    civic + memhobby + fq5read + frlxnatr + lonely  + sfintfpn+
    op5happe + opwslpw + op3ecn3a + op5chnca + axecnsf,  data=js_
    data1b)
summary(olsb1m)
#Adjusted Rquare higher
coeftest(olsb1m, df = Inf, vcov = vcovHC(olsb1m, type="HC"))
#Age now least significant
olsb1n <- lm(st5job ~ agebSquare + male + never_married  +
    graduatedCollege + szincomx + xjobhwkLog + executive  +
    dispatched  + family_worker + szcmttl + st5hlthy  + st5leisy +
    civic + memhobby + fq5read + frlxnatr + lonely  + sfintfpn+
    op5happe + opwslpw + op3ecn3a + op5chnca + axecnsf,  data=js_
    data1b)
summary(olsb1n)
#Adjusted Rquare higher
coeftest(olsb1n, df = Inf, vcov = vcovHC(olsb1n, type="HC"))
#Satisfaction with health state least significant
olsb1o <- lm(st5job ~ agebSquare + male + never_married  +
    graduatedCollege + szincomx + xjobhwkLog + executive  +
    dispatched  + family_worker + szcmttl + st5leisy + civic +
    memhobby + fq5read + frlxnatr + lonely  + sfintfpn+ op5happe +
```

```
    opwslpw + op3ecn3a + op5chnca + axecnsf,  data=js_data1b)
summary(olsb1o)
#Adjusted Rquare same
coeftest(olsb1o, df = Inf, vcov = vcovHC(olsb1o, type="HC"))
#Membership in civil/social organization now least significant
olsb1p <- lm(st5job ~ agebSquare + male + never_married  +
    graduatedCollege + szincomx + xjobhwkLog + executive  +
    dispatched  + family_worker + szcmttl + st5leisy  + memhobby +
    fq5read + frlxnatr + lonely  + sfintfpn+ op5happe + opwslpw +
    op3ecn3a + op5chnca + axecnsf,  data=js_data1b)
summary(olsb1p)
#Adjusted Rquare higher
coeftest(olsb1p, df = Inf, vcov = vcovHC(olsb1p, type="HC"))
#Time in nature now least significant
olsb1q <- lm(st5job ~ agebSquare + male + never_married  +
    graduatedCollege + szincomx + xjobhwkLog + executive  +
    dispatched  + family_worker + szcmttl + st5leisy  + memhobby +
    fq5read  + lonely  + sfintfpn+ op5happe + opwslpw + op3ecn3a +
    op5chnca + axecnsf,  data=js_data1b)
summary(olsb1q)
#Adjusted Rquare same
coeftest(olsb1q, df = Inf, vcov = vcovHC(olsb1q, type="HC"))
#Mmbership in hobby organization now least significant
olsb1r <- lm(st5job ~ agebSquare + male + never_married  +
    graduatedCollege + szincomx + xjobhwkLog + executive  +
    dispatched  + family_worker + szcmttl + st5leisy + fq5read  +
    lonely  + sfintfpn+ op5happe + opwslpw + op3ecn3a + op5chnca +
    axecnsf,  data=js_data1b)
summary(olsb1r)
#Adjusted Rquare same
coeftest(olsb1r, df = Inf, vcov = vcovHC(olsb1r, type="HC"))
#Commuting time now least significant
olsb1s <- lm(st5job ~ agebSquare + male + never_married  +
    graduatedCollege + szincomx + xjobhwkLog + executive  +
    dispatched  + family_worker + st5leisy + fq5read  + lonely  +
    sfintfpn+ op5happe + opwslpw + op3ecn3a + op5chnca + axecnsf,
    data=js_data1b)
summary(olsb1s)
#Stopping at previous model (olsb1r) since adjusted r square for
    olsb1s is lower

#Descriptive Statistics for general model of subsample B
summary(js_data1b$st5job)
se(js_data1b$st5job)
summary(js_data1b$ageb)
se(js_data1b$ageb)
summary(js_data1b$male)
se(js_data1b$male)
summary(js_data1b$never_married)
se(js_data1b$never_married)
summary(js_data1b$divorced)
se(js_data1b$divorced)
summary(js_data1b$widowed)
```

```
se(js_data1b$widowed)
summary(js_data1b$ccnumttl)
se(js_data1b$ccnumttl)
summary(js_data1b$graduatedCollege)
se(js_data1b$graduatedCollege)
summary(js_data1b$szincomx)
se(js_data1b$szincomx)
summary(js_data1b$xjobhwk)
se(js_data1b$xjobhwk)
summary(js_data1b$executive)
se(js_data1b$executive)
summary(js_data1b$temporary)
se(js_data1b$temporary)
summary(js_data1b$dispatched)
se(js_data1b$dispatched)
summary(js_data1b$self_employed)
se(js_data1b$self_employed)
summary(js_data1b$family_worker)
se(js_data1b$family_worker)
summary(js_data1b$GovermentAgency)
se(js_data1b$GovermentAgency)
summary(js_data1b$szcmttl)
se(js_data1b$szcmttl)
summary(js_data1b$st5hlthy)
se(js_data1b$st5hlthy)
summary(js_data1b$st5areay)
se(js_data1b$st5areay)
summary(js_data1b$st5leisy)
se(js_data1b$st5leisy)
summary(js_data1b$civic)
se(js_data1b$civic)
summary(js_data1b$memhobby)
se(js_data1b$memhobby)
summary(js_data1b$fq7frsee)
se(js_data1b$fq7frsee)
summary(js_data1b$fqsport)
se(js_data1b$fqsport)
summary(js_data1b$fq5read)
se(js_data1b$fq5read)
summary(js_data1b$hrtv)
se(js_data1b$hrtv)
summary(js_data1b$lonely)
se(js_data1b$lonely)
summary(js_data1b$sfintfpn)
se(js_data1b$sfintfpn)
summary(js_data1b$op5happe)
se(js_data1b$op5happe)
summary(js_data1b$opwslpw)
se(js_data1b$opwslpw)
summary(js_data1b$op3ecn3a)
se(js_data1b$op3ecn3a)
summary(js_data1b$op5chnca)
se(js_data1b$op5chnca)
```

```
summary(js_data1b$op5levk)
se(js_data1b$op5levk)
summary(js_data1b$axecnsf)
se(js_data1b$axecnsf)


#Estimation of regression for famale sample
js_data1bFem <- js_data1b
#Ereasing male sample
js_data1bFem <- subset(js_data1bFem,js_data1bFem$sexa != 1)
#Estimation OLS
olsb1Fem <- lm(st5job ~ ageb + agebSquare+ never_married + divorced +
     widowed + ccnumttl + graduatedCollege + szincomx + xjobhwkLog +
     executive + temporary + dispatched + self_employed + family_
     worker + GovermentAgency + szcmttl + st5hlthy + st5areay +
     st5leisy + civic + memhobby + fq7frsee + fqsport + fq5read +
     frlxnatr + lonely + op4trust + sfintfpn + sfhlcnd + op5happe +
     opwslpw + op3ecn3a + op5chnca + op5levk + axecnsf,  js_data1bFem)
summary(olsb1Fem)
coeftest(olsb1Fem, df = Inf, vcov = vcovHC(olsb1Fem, type="HC"))
#Being divorced least significant
olsb1Fema <- lm(st5job ~ ageb + agebSquare+ never_married + widowed +
     ccnumttl + graduatedCollege + szincomx + xjobhwkLog + executive
     + temporary + dispatched + self_employed + family_worker +
     GovermentAgency + szcmttl + st5hlthy + st5areay + st5leisy +
     civic + memhobby + fq7frsee + fqsport + fq5read + frlxnatr +
     lonely + op4trust + sfintfpn + sfhlcnd + op5happe + opwslpw +
     op3ecn3a + op5chnca + op5levk + axecnsf,  js_data1bFem)
summary(olsb1Fema)
#Adjusted Rquare higher
coeftest(olsb1Fema, df = Inf, vcov = vcovHC(olsb1Fema, type="HC"))
#Satisfaction with area of living now least significant
olsb1Femb <- lm(st5job ~ ageb + agebSquare+ never_married + widowed +
     ccnumttl + graduatedCollege + szincomx + xjobhwkLog + executive
     + temporary + dispatched + self_employed + family_worker +
     GovermentAgency + szcmttl + st5hlthy + st5leisy + civic +
     memhobby + fq7frsee + fqsport + fq5read + frlxnatr + lonely +
     op4trust + sfintfpn + sfhlcnd + op5happe + opwslpw + op3ecn3a +
     op5chnca + op5levk + axecnsf,  js_data1bFem)
summary(olsb1Femb)
#Adjusted Rquare higher
coeftest(olsb1Femb, df = Inf, vcov = vcovHC(olsb1Femb, type="HC"))
#Working in Government Agency now least significant
olsb1Femc <- lm(st5job ~ ageb + agebSquare+ never_married + widowed +
     ccnumttl + graduatedCollege + szincomx + xjobhwkLog + executive
     + temporary + dispatched + self_employed + family_worker  +
     szcmttl + st5hlthy + st5leisy + civic + memhobby + fq7frsee +
     fqsport + fq5read + frlxnatr + lonely + op4trust + sfintfpn +
     sfhlcnd + op5happe + opwslpw + op3ecn3a + op5chnca + op5levk +
     axecnsf,  js_data1bFem)
summary(olsb1Femc)
#Adjusted Rquare higher
coeftest(olsb1Femc, df = Inf, vcov = vcovHC(olsb1Femc, type="HC"))
```

```
#Time nature now least significant
olsb1Femd <- lm(st5job ~ ageb + agebSquare+ never_married + widowed +
      ccnumttl + graduatedCollege + szincomx + xjobhwkLog + executive
    + temporary + dispatched + self_employed + family_worker  +
    szcmttl + st5hlthy + st5leisy + civic + memhobby + fq7frsee +
    fqsport + fq5read + lonely + op4trust + sfintfpn + sfhlcnd +
    op5happe + opwslpw + op3ecn3a + op5chnca + op5levk + axecnsf,  js
    _data1bFem)
summary(olsb1Femd)
#Adjusted Rquare higher
coeftest(olsb1Femd, df = Inf, vcov = vcovHC(olsb1Femd, type="HC"))
#Number of books now least significant
olsb1Feme <- lm(st5job ~ ageb + agebSquare+ never_married + widowed +
      ccnumttl + graduatedCollege + szincomx + xjobhwkLog + executive
    + temporary + dispatched + self_employed + family_worker  +
    szcmttl + st5hlthy + st5leisy + civic + memhobby + fq7frsee +
    fqsport  + lonely + op4trust + sfintfpn + sfhlcnd + op5happe +
    opwslpw + op3ecn3a + op5chnca + op5levk + axecnsf,  js_data1bFem)
summary(olsb1Feme)
#Adjusted Rquare higher
coeftest(olsb1Feme, df = Inf, vcov = vcovHC(olsb1Feme, type="HC"))
#Being Family worker now least significant
olsb1Femf <- lm(st5job ~ ageb + agebSquare+ never_married + widowed +
      ccnumttl + graduatedCollege + szincomx + xjobhwkLog + executive
    + temporary + dispatched + self_employed + szcmttl + st5hlthy +
    st5leisy + civic + memhobby + fq7frsee + fqsport  + lonely +
    op4trust + sfintfpn + sfhlcnd + op5happe + opwslpw + op3ecn3a +
    op5chnca + op5levk + axecnsf,  js_data1bFem)
summary(olsb1Femf)
#Adjusted Rquare higher
coeftest(olsb1Femf, df = Inf, vcov = vcovHC(olsb1Femf, type="HC"))
#Satisfaction with leisure now leastt significant
olsb1Femg <- lm(st5job ~ ageb + agebSquare+ never_married + widowed +
      ccnumttl + graduatedCollege + szincomx + xjobhwkLog + executive
    + temporary + dispatched + self_employed + szcmttl + st5hlthy  +
    civic + memhobby + fq7frsee + fqsport  + lonely + op4trust +
    sfintfpn + sfhlcnd + op5happe + opwslpw + op3ecn3a + op5chnca +
    op5levk + axecnsf,  js_data1bFem)
summary(olsb1Femg)
#Adjusted Rquare higher
coeftest(olsb1Femg, df = Inf, vcov = vcovHC(olsb1Femg, type="HC"))
#Number of children now least significant
olsb1Femh <- lm(st5job ~ ageb + agebSquare+ never_married + widowed +
      graduatedCollege + szincomx + xjobhwkLog + executive + temporary
      + dispatched + self_employed + szcmttl + st5hlthy  + civic +
    memhobby + fq7frsee + fqsport  + lonely + op4trust + sfintfpn +
    sfhlcnd + op5happe + opwslpw + op3ecn3a + op5chnca + op5levk +
    axecnsf,  js_data1bFem)
summary(olsb1Femh)
#Adjusted Rquare higher
coeftest(olsb1Femh, df = Inf, vcov = vcovHC(olsb1Femh, type="HC"))
#Membership in civil/civic organization now least signficant
```

```
olsb1Femi <- lm(st5job ~ ageb + agebSquare+ never_married + widowed +
     graduatedCollege + szincomx + xjobhwkLog + executive + temporary
     + dispatched + self_employed + szcmttl + st5hlthy + memhobby +
    fq7frsee + fqsport  + lonely + op4trust + sfintfpn + sfhlcnd +
    op5happe + opwslpw + op3ecn3a + op5chnca + op5levk + axecnsf ,  js
    _data1bFem)
summary(olsb1Femi)
#Adjusted Rquare higher
coeftest(olsb1Femi, df = Inf, vcov = vcovHC(olsb1Femi, type="HC"))
#Self-reported health condition now least significant
olsb1Femj <- lm(st5job ~ ageb + agebSquare+ never_married + widowed +
     graduatedCollege + szincomx + xjobhwkLog + executive + temporary
     + dispatched + self_employed + szcmttl + st5hlthy + memhobby +
    fq7frsee + fqsport  + lonely + op4trust + sfintfpn  + op5happe +
    opwslpw + op3ecn3a + op5chnca + op5levk + axecnsf ,  js_data1bFem)
summary(olsb1Femj)
#Adjusted Rquare higher
coeftest(olsb1Femj, df = Inf, vcov = vcovHC(olsb1Femj, type="HC"))
#Anxiety in regard to future economic situtuation least significant
olsb1Femk <- lm(st5job ~ ageb + agebSquare+ never_married + widowed +
     graduatedCollege + szincomx + xjobhwkLog + executive + temporary
     + dispatched + self_employed + szcmttl + st5hlthy + memhobby +
    fq7frsee + fqsport  + lonely + op4trust + sfintfpn  + op5happe +
    opwslpw + op3ecn3a + op5chnca + op5levk ,  js_data1bFem)
summary(olsb1Femk)
#Adjusted Rquare higher
coeftest(olsb1Femk, df = Inf, vcov = vcovHC(olsb1Femk, type="HC"))
#Pain now least significant
olsb1Feml <- lm(st5job ~ ageb + agebSquare+ never_married + widowed +
     graduatedCollege + szincomx + xjobhwkLog + executive + temporary
     + dispatched + self_employed + szcmttl + st5hlthy + memhobby +
    fq7frsee + fqsport  + lonely + op4trust + op5happe + opwslpw +
    op3ecn3a + op5chnca + op5levk ,  js_data1bFem)
summary(olsb1Feml)
#Adjusted Rquare higher
coeftest(olsb1Feml, df = Inf, vcov = vcovHC(olsb1Feml, type="HC"))
#Executive position now least signficant
olsb1Femm <- lm(st5job ~ ageb + agebSquare+ never_married + widowed +
     graduatedCollege + szincomx + xjobhwkLog  + temporary +
    dispatched + self_employed + szcmttl + st5hlthy + memhobby +
    fq7frsee + fqsport  + lonely + op4trust + op5happe + opwslpw +
    op3ecn3a + op5chnca + op5levk ,  js_data1bFem)
summary(olsb1Femm)
#Adjusted Rquare higher
coeftest(olsb1Femm, df = Inf, vcov = vcovHC(olsb1Femm, type="HC"))
#Trust now least significant
olsb1Femn <- lm(st5job ~ ageb + agebSquare+ never_married + widowed +
     graduatedCollege + szincomx + xjobhwkLog  + temporary +
    dispatched + self_employed + szcmttl + st5hlthy + memhobby +
    fq7frsee + fqsport  + lonely + op5happe + opwslpw + op3ecn3a +
    op5chnca + op5levk ,  js_data1bFem)
summary(olsb1Femn)
#Adjusted Rquare higher
```

```
coeftest(olsb1Femn, df = Inf, vcov = vcovHC(olsb1Femn, type="HC"))
#Self reported social status now least significant
olsb1Femo <- lm(st5job ~ ageb + agebSquare+ never_married + widowed + graduatedCollege + szincomx + xjobhwkLog + temporary + dispatched + self_employed + szcmttl + st5hlthy + memhobby + fq7frsee + fqsport + lonely + op5happe + opwslpw + op3ecn3a + op5chnca , js_data1bFem)
summary(olsb1Femo)
#Adjusted Rquare higher
coeftest(olsb1Femo, df = Inf, vcov = vcovHC(olsb1Femo, type="HC"))
#oncome now least significant
olsb1Femp <- lm(st5job ~ ageb + agebSquare+ never_married + widowed + graduatedCollege + xjobhwkLog + temporary + dispatched + self_employed + szcmttl + st5hlthy + memhobby + fq7frsee + fqsport + lonely + op5happe + opwslpw + op3ecn3a + op5chnca , js_data1bFem)
summary(olsb1Femp)
#Adjusted Rquare higher
coeftest(olsb1Femp, df = Inf, vcov = vcovHC(olsb1Femp, type="HC"))
#Doing sports now least significant
olsb1Femq <- lm(st5job ~ ageb + agebSquare+ never_married + widowed + graduatedCollege + xjobhwkLog + temporary + dispatched + self_employed + szcmttl + st5hlthy + memhobby + fq7frsee + lonely + op5happe + opwslpw + op3ecn3a + op5chnca , js_data1bFem)
summary(olsb1Femq)
#Adjusted Rquare higher
coeftest(olsb1Femq, df = Inf, vcov = vcovHC(olsb1Femq, type="HC"))
#Being dispatched worker now least significant
olsb1Femr <- lm(st5job ~ ageb + agebSquare+ never_married + widowed + graduatedCollege + xjobhwkLog + temporary + self_employed + szcmttl + st5hlthy + memhobby + fq7frsee + lonely + op5happe + opwslpw + op3ecn3a + op5chnca , js_data1bFem)
summary(olsb1Femr)
#Adjusted Rquare higher
coeftest(olsb1Femr, df = Inf, vcov = vcovHC(olsb1Femr, type="HC"))
#Membership in hobby organzation now least significant
olsb1Fems <- lm(st5job ~ ageb + agebSquare+ never_married + widowed + graduatedCollege + xjobhwkLog + temporary + self_employed + szcmttl + st5hlthy + fq7frsee + lonely + op5happe + opwslpw + op3ecn3a + op5chnca , js_data1bFem)
summary(olsb1Fems)
#Adjusted Rquare higher
coeftest(olsb1Fems, df = Inf, vcov = vcovHC(olsb1Fems, type="HC"))
#Never-married now least significant
olsb1Femt <- lm(st5job ~ ageb + agebSquare + widowed + graduatedCollege + xjobhwkLog + temporary + self_employed + szcmttl + st5hlthy + fq7frsee + lonely + op5happe + opwslpw + op3ecn3a + op5chnca , js_data1bFem)
summary(olsb1Femt)
#Adjusted Rquare higher
coeftest(olsb1Femt, df = Inf, vcov = vcovHC(olsb1Femt, type="HC"))
#Age now least significant
```

```
olsb1Femu <- lm(st5job ~ agebSquare + widowed + graduatedCollege + xjobhwkLog + temporary + self_employed + szcmttl + st5hlthy + fq7frsee + lonely + op5happe + opwslpw + op3ecn3a + op5chnca , js_data1bFem)
summary(olsb1Femu)
#Adjusted Rquare higher
coeftest(olsb1Femu, df = Inf, vcov = vcovHC(olsb1Femu, type="HC"))
#Widowed now least significant
olsb1Femv <- lm(st5job ~ agebSquare + graduatedCollege + xjobhwkLog + temporary + self_employed + szcmttl + st5hlthy + fq7frsee + lonely + op5happe + opwslpw + op3ecn3a + op5chnca , js_data1bFem )
summary(olsb1Femv)
#Adjusted Rquare higher
coeftest(olsb1Femv, df = Inf, vcov = vcovHC(olsb1Femv, type="HC"))
#Working hours now least significant
olsb1Femw <- lm(st5job ~ agebSquare + graduatedCollege + temporary + self_employed + szcmttl + st5hlthy + fq7frsee + lonely + op5happe + opwslpw + op3ecn3a + op5chnca , js_data1bFem)
summary(olsb1Femw)
#Adjusted Rquare higher
coeftest(olsb1Femw, df = Inf, vcov = vcovHC(olsb1Femw, type="HC"))
#Working temporary now least signifcant
olsb1Femx <- lm(st5job ~ agebSquare + graduatedCollege + self_employed + szcmttl + st5hlthy + fq7frsee + lonely + op5happe + opwslpw + op3ecn3a + op5chnca , js_data1bFem)
summary(olsb1Femx)
#Adjusted Rquare lower, but seeing friends has a similar p value
olsb1Femx <- lm(st5job ~ agebSquare + graduatedCollege + temporary + self_employed + szcmttl + st5hlthy + lonely + op5happe + opwslpw + op3ecn3a + op5chnca , js_data1bFem)
summary(olsb1Femx)
#Stopping at previous model (olsb1Femw) since adjusted r square for both olsb1Femx is lower

#Descriptive Statistics for female model of subsample B
summary(js_data1bFem$st5job)
se(js_data1bFem$st5job)
summary(js_data1bFem$ageb)
se(js_data1bFem$ageb)
summary(js_data1bFem$male)
se(js_data1bFem$male)
summary(js_data1bFem$never_married)
se(js_data1bFem$never_married)
summary(js_data1bFem$divorced)
se(js_data1bFem$divorced)
summary(js_data1bFem$widowed)
se(js_data1bFem$widowed)
summary(js_data1bFem$ccnumttl)
se(js_data1bFem$ccnumttl)
summary(js_data1bFem$graduatedCollege)
se(js_data1bFem$graduatedCollege)
summary(js_data1bFem$szincomx)
```

```
se(js_data1bFem$szincomx)
summary(js_data1bFem$xjobhwk)
se(js_data1bFem$xjobhwk)
summary(js_data1bFem$executive)
se(js_data1bFem$executive)
summary(js_data1bFem$temporary)
se(js_data1bFem$temporary)
summary(js_data1bFem$dispatched)
se(js_data1bFem$dispatched)
summary(js_data1bFem$self_employed)
se(js_data1bFem$self_employed)
summary(js_data1bFem$family_worker)
se(js_data1bFem$family_worker)
summary(js_data1bFem$GovermentAgency)
se(js_data1bFem$GovermentAgency)
summary(js_data1bFem$szcmttl)
se(js_data1bFem$szcmttl)
summary(js_data1bFem$st5hlthy)
se(js_data1bFem$st5hlthy)
summary(js_data1bFem$st5areay)
se(js_data1bFem$st5areay)
summary(js_data1bFem$st5leisy)
se(js_data1bFem$st5leisy)
summary(js_data1bFem$civic)
se(js_data1bFem$civic)
summary(js_data1bFem$memhobby)
se(js_data1bFem$memhobby)
summary(js_data1bFem$fq7frsee)
se(js_data1bFem$fq7frsee)
summary(js_data1bFem$fqsport)
se(js_data1bFem$fqsport)
summary(js_data1bFem$fq5read)
se(js_data1bFem$fq5read)
summary(js_data1bFem$hrtv)
se(js_data1bFem$hrtv)
summary(js_data1bFem$lonely)
se(js_data1bFem$lonely)
summary(js_data1bFem$sfintfpn)
se(js_data1bFem$sfintfpn)
summary(js_data1bFem$op5happe)
se(js_data1bFem$op5happe)
summary(js_data1bFem$opwslpw)
se(js_data1bFem$opwslpw)
summary(js_data1bFem$op3ecn3a)
se(js_data1bFem$op3ecn3a)
summary(js_data1bFem$op5chnca)
se(js_data1bFem$op5chnca)
summary(js_data1bFem$op5levk)
se(js_data1bFem$op5levk)
summary(js_data1bFem$axecnsf)
se(js_data1bFem$axecnsf)
```

```
#Estimation of regression for male sample
js_data1bMal <- js_data1b
#Ereasing male sample
js_data1bMal <- subset(js_data1bMal,js_data1bMal$sexa != 0)
#Estimation OLS
olsb1Mal <- lm(st5job ~ ageb + agebSquare+ never_married + divorced +
    widowed + ccnumttl + graduatedCollege + szincomx + xjobhwkLog +
    executive + temporary + dispatched + self_employed + family_
    worker + GovermentAgency + szcmttl + st5hlthy + st5areay +
    st5leisy + civic + memhobby + fq7frsee + fqsport + fq5read +
    frlxnatr + lonely + op4trust + sfintfpn + sfhlcnd + op5happe +
    opwslpw + op3ecn3a + op5chnca + op5levk + axecnsf , js_data1bMal)
summary(olsb1Mal)
coeftest(olsb1Mal, df = Inf, vcov = vcovHC(olsb1Mal, type="HC"))
#Commuting time least significant
olsb1Mala <- lm(st5job ~ ageb + agebSquare+ never_married + divorced
    + widowed + ccnumttl + graduatedCollege + szincomx + xjobhwkLog +
    executive + temporary + dispatched + self_employed + family_
    worker + GovermentAgency + st5hlthy + st5areay + st5leisy + civic
    + memhobby + fq7frsee + fqsport + fq5read + frlxnatr + lonely +
    op4trust + sfintfpn + sfhlcnd + op5happe + opwslpw + op3ecn3a +
    op5chnca + op5levk + axecnsf , js_data1bMal)
summary(olsb1Mala)
#Adjusted Rquare higher
coeftest(olsb1Mala, df = Inf, vcov = vcovHC(olsb1Mala, type="HC"))
#Seeing friends now least significant
olsb1Malb <- lm(st5job ~ ageb + agebSquare+ never_married + divorced
    + widowed + ccnumttl + graduatedCollege + szincomx + xjobhwkLog +
    executive + temporary + dispatched + self_employed + family_
    worker + GovermentAgency + st5hlthy + st5areay + st5leisy + civic
    + memhobby  + fqsport + fq5read + frlxnatr + lonely + op4trust +
    sfintfpn + sfhlcnd + op5happe + opwslpw + op3ecn3a + op5chnca +
    op5levk + axecnsf , js_data1bMal)
summary(olsb1Malb)
#Adjusted Rquare higher
coeftest(olsb1Malb, df = Inf, vcov = vcovHC(olsb1Malb, type="HC"))
#Working time now least significant
olsb1Malc <- lm(st5job ~ ageb + agebSquare+ never_married + divorced
    + widowed + ccnumttl + graduatedCollege + szincomx + executive +
    temporary + dispatched + self_employed + family_worker +
    GovermentAgency + st5hlthy + st5areay + st5leisy + civic +
    memhobby  + fqsport + fq5read + frlxnatr + lonely + op4trust +
    sfintfpn + sfhlcnd + op5happe + opwslpw + op3ecn3a + op5chnca +
    op5levk + axecnsf , js_data1bMal)
summary(olsb1Malc)
#Adjusted Rquare higher
coeftest(olsb1Malc, df = Inf, vcov = vcovHC(olsb1Malc, type="HC"))
#Doing sports now least significant
olsb1Mald <- lm(st5job ~ ageb + agebSquare+ never_married + divorced
    + widowed + ccnumttl + szincomx + executive + temporary +
    dispatched + self_employed + family_worker + GovermentAgency +
    st5hlthy + st5areay + st5leisy + civic + memhobby  + fqsport +
    fq5read + frlxnatr + lonely + op4trust + sfintfpn + sfhlcnd +
```

```r
    op5happe + opwslpw + op3ecn3a + op5chnca + op5levk + axecnsf, js_
    data1bMal)
summary(olsb1Mald)
#Adjusted Rquare higher
coeftest(olsb1Mald, df = Inf, vcov = vcovHC(olsb1Mald, type="HC"))
#Being temporary worker now least significant
olsb1Male <- lm(st5job ~ ageb + agebSquare+ never_married + divorced
    + widowed + ccnumttl + szincomx + executive + dispatched + self_
    employed + family_worker + GovermentAgency + st5hlthy + st5areay
    + st5leisy + civic + memhobby  + fqsport + fq5read + frlxnatr +
    lonely + op4trust + sfintfpn + sfhlcnd + op5happe + opwslpw +
    op3ecn3a + op5chnca + op5levk + axecnsf, js_data1bMal)
summary(olsb1Male)
#Adjusted Rquare higher
coeftest(olsb1Male, df = Inf, vcov = vcovHC(olsb1Male, type="HC"))
#Doing sports now least significant
olsb1Malf <- lm(st5job ~ ageb + agebSquare+ never_married + divorced
    + widowed + ccnumttl + szincomx + executive + dispatched + self_
    employed + family_worker + GovermentAgency + st5hlthy + st5areay
    + st5leisy + civic + memhobby  + fq5read + frlxnatr + lonely +
    op4trust + sfintfpn + sfhlcnd + op5happe + opwslpw + op3ecn3a +
    op5chnca + op5levk + axecnsf, js_data1bMal)
summary(olsb1Malf)
#Adjusted Rquare higher
coeftest(olsb1Malf, df = Inf, vcov = vcovHC(olsb1Malf, type="HC"))
#Satisfaction with living are now least significant
olsb1Malg <- lm(st5job ~ ageb + agebSquare+ never_married + divorced
    + widowed + ccnumttl + szincomx + executive + dispatched + self_
    employed + family_worker + GovermentAgency + st5hlthy  + st5leisy
     + civic + memhobby  + fq5read + frlxnatr + lonely + op4trust +
    sfintfpn + sfhlcnd + op5happe + opwslpw + op3ecn3a + op5chnca +
    op5levk + axecnsf, js_data1bMal)
summary(olsb1Malg)
#Adjusted Rquare higher
coeftest(olsb1Malg, df = Inf, vcov = vcovHC(olsb1Malg, type="HC"))
#Self-reported health condition now least significant
olsb1Malh <- lm(st5job ~ ageb + agebSquare+ never_married + divorced
    + widowed + ccnumttl + szincomx + executive + dispatched + self_
    employed + family_worker + GovermentAgency + st5hlthy  + st5leisy
     + civic + memhobby  + fq5read + frlxnatr + lonely + op4trust +
    sfintfpn  + op5happe + opwslpw + op3ecn3a + op5chnca + op5levk +
    axecnsf, js_data1bMal)
summary(olsb1Malh)
#Adjusted Rquare higher
coeftest(olsb1Malh, df = Inf, vcov = vcovHC(olsb1Malh, type="HC"))
#Divorced now least signifcant
olsb1Mali <- lm(st5job ~ ageb + agebSquare+ never_married + widowed +
     ccnumttl + szincomx + executive + dispatched + self_employed +
    family_worker + GovermentAgency + st5hlthy  + st5leisy + civic +
    memhobby  + fq5read + frlxnatr + lonely + op4trust + sfintfpn  +
    op5happe + opwslpw + op3ecn3a + op5chnca + op5levk + axecnsf, js_
    data1bMal)
summary(olsb1Mali)
```

```
#Adjusted Rquare higher
coeftest(olsb1Mali, df = Inf, vcov = vcovHC(olsb1Mali, type="HC"))
#Number of children now least significant
olsb1Malj <- lm(st5job ~ ageb + agebSquare+ never_married + widowed + szincomx + executive + dispatched + self_employed + family_worker + GovermentAgency + st5hlthy  + st5leisy + civic + memhobby  + fq5read + frlxnatr + lonely + op4trust + sfintfpn  + op5happe + opwslpw + op3ecn3a + op5chnca + op5levk + axecnsf, js_data1bMal)
summary(olsb1Malj)
#Adjusted Rquare higher
coeftest(olsb1Malj, df = Inf, vcov = vcovHC(olsb1Malj, type="HC"))
#Trust now least signifcant
olsb1Malk <- lm(st5job ~ ageb + agebSquare+ never_married + widowed + szincomx + executive + dispatched + self_employed + family_worker + GovermentAgency + st5hlthy  + st5leisy + civic + memhobby  + fq5read + frlxnatr + lonely  + sfintfpn  + op5happe + opwslpw + op3ecn3a + op5chnca + op5levk + axecnsf, js_data1bMal)
summary(olsb1Malk)
#Adjusted Rquare higher
coeftest(olsb1Malk, df = Inf, vcov = vcovHC(olsb1Malk, type="HC"))
#Feeling lonely now least significant
olsb1Mall <- lm(st5job ~ ageb + agebSquare+ never_married + widowed + szincomx + executive + dispatched + self_employed + family_worker + GovermentAgency + st5hlthy  + st5leisy + civic + memhobby  + fq5read + frlxnatr + sfintfpn  + op5happe + opwslpw + op3ecn3a + op5chnca + op5levk + axecnsf, js_data1bMal)
summary(olsb1Mall)
#Adjusted Rquare higher
coeftest(olsb1Mall, df = Inf, vcov = vcovHC(olsb1Mall, type="HC"))
#Satisfaction with health condition now least signifcant
olsb1Malm <- lm(st5job ~ ageb + agebSquare+ never_married + widowed + szincomx + executive + dispatched + self_employed + family_worker + GovermentAgency   + st5leisy + civic + memhobby  + fq5read + frlxnatr + sfintfpn  + op5happe + opwslpw + op3ecn3a + op5chnca + op5levk + axecnsf, js_data1bMal)
summary(olsb1Malm)
#Adjusted Rquare higher
coeftest(olsb1Malm, df = Inf, vcov = vcovHC(olsb1Malm, type="HC"))
#Age now least significant
olsb1Maln <- lm(st5job ~  agebSquare+ never_married + widowed  + szincomx + executive + dispatched + self_employed + family_worker + GovermentAgency   + st5leisy + civic + memhobby  + fq5read + frlxnatr + sfintfpn  + op5happe + opwslpw + op3ecn3a + op5chnca + op5levk + axecnsf, js_data1bMal)
summary(olsb1Maln)
#Adjusted Rquare higher
coeftest(olsb1Maln, df = Inf, vcov = vcovHC(olsb1Maln, type="HC"))
#Working in Government oragnization now least significant
olsb1Malo <- lm(st5job ~  agebSquare+ never_married + widowed  + szincomx + executive + dispatched + self_employed + family_worker  + st5leisy + civic + memhobby  + fq5read + frlxnatr + sfintfpn  + op5happe + opwslpw + op3ecn3a + op5chnca + op5levk + axecnsf,
```

```
js_data1bMal)
summary(olsb1Malo)
#Adjusted Rquare higher
coeftest(olsb1Malo, df = Inf, vcov = vcovHC(olsb1Malo, type="HC"))
#Time in mature now least significant
olsb1Malp <- lm(st5job ~ agebSquare+ never_married + widowed + szincomx + executive + dispatched + self_employed + family_worker + st5leisy + civic + memhobby + fq5read + sfintfpn + op5happe + opwslpw + op3ecn3a + op5chnca + op5levk + axecnsf, js_data1bMal)
summary(olsb1Malp)
#Adjusted Rquare higher
coeftest(olsb1Malp, df = Inf, vcov = vcovHC(olsb1Malp, type="HC"))
#Membership in hobby organization now least significant
olsb1Malq <- lm(st5job ~ agebSquare+ never_married + widowed + szincomx + executive + dispatched + self_employed + family_worker + st5leisy + civic + fq5read + sfintfpn + op5happe + opwslpw + op3ecn3a + op5chnca + op5levk + axecnsf, js_data1bMal)
summary(olsb1Malq)
#Adjusted Rquare higher
coeftest(olsb1Malq, df = Inf, vcov = vcovHC(olsb1Malq, type="HC"))
#Satisfaction lesure now least significant
olsb1Malr <- lm(st5job ~ agebSquare+ never_married + widowed + szincomx + executive + dispatched + self_employed + family_worker + civic + fq5read + sfintfpn + op5happe + opwslpw + op3ecn3a + op5chnca + op5levk + axecnsf, js_data1bMal)
summary(olsb1Malr)
#Stopping at previous model (olsb1Malq) since adjusted r square for olsb1Malr is lower

#Testing influence of other variables ond leisure satisfaction variable
olsb1Malqb <- lm(st5job ~ agebSquare+ never_married + widowed + szincomx + executive + dispatched + self_employed + family_worker + st5leisy + civic + fq5read + opwslpw + op3ecn3a + op5chnca + axecnsf + op5levk, js_data1bMal)
summary(olsb1Malqb)
#removel of happiness and pain makes satisfaction with leisure significant again.

#Descriptive Statistics for female model of subsample B
summary(js_data1bMal$st5job)
se(js_data1bMal$st5job)
summary(js_data1bMal$ageb)
se(js_data1bMal$ageb)
summary(js_data1bMal$male)
se(js_data1bMal$male)
summary(js_data1bMal$never_married)
se(js_data1bMal$never_married)
summary(js_data1bMal$divorced)
se(js_data1bMal$divorced)
summary(js_data1bMal$widowed)
se(js_data1bMal$widowed)
```

```
summary(js_data1bMal$ccnumttl)
se(js_data1bMal$ccnumttl)
summary(js_data1bMal$graduatedCollege)
se(js_data1bMal$graduatedCollege)
summary(js_data1bMal$szincomx)
se(js_data1bMal$szincomx)
summary(js_data1bMal$xjobhwk)
se(js_data1bMal$xjobhwk)
summary(js_data1bMal$executive)
se(js_data1bMal$executive)
summary(js_data1bMal$temporary)
se(js_data1bMal$temporary)
summary(js_data1bMal$dispatched)
se(js_data1bMal$dispatched)
summary(js_data1bMal$self_employed)
se(js_data1bMal$self_employed)
summary(js_data1bMal$family_worker)
se(js_data1bMal$family_worker)
summary(js_data1bMal$GovermentAgency)
se(js_data1bMal$GovermentAgency)
summary(js_data1bMal$szcmttl)
se(js_data1bMal$szcmttl)
summary(js_data1bMal$st5hlthy)
se(js_data1bMal$st5hlthy)
summary(js_data1bMal$st5areay)
se(js_data1bMal$st5areay)
summary(js_data1bMal$st5leisy)
se(js_data1bMal$st5leisy)
summary(js_data1bMal$civic)
se(js_data1bMal$civic)
summary(js_data1bMal$memhobby)
se(js_data1bMal$memhobby)
summary(js_data1bMal$fq7frsee)
se(js_data1bMal$fq7frsee)
summary(js_data1bMal$fqsport)
se(js_data1bMal$fqsport)
summary(js_data1bMal$fq5read)
se(js_data1bMal$fq5read)
summary(js_data1bMal$hrtv)
se(js_data1bMal$hrtv)
summary(js_data1bMal$lonely)
se(js_data1bMal$lonely)
summary(js_data1bMal$sfintfpn)
se(js_data1bMal$sfintfpn)
summary(js_data1bMal$op5happe)
se(js_data1bMal$op5happe)
summary(js_data1bMal$opwslpw)
se(js_data1bMal$opwslpw)
summary(js_data1bMal$op3ecn3a)
se(js_data1bMal$op3ecn3a)
summary(js_data1bMal$op5chnca)
se(js_data1bMal$op5chnca)
summary(js_data1bMal$op5levk)
```

```r
se(js_data1bMal$op5levk)
summary(js_data1bMal$axecnsf)
se(js_data1bMal$axecnsf)

stargazer(list(coeftest(olsb1r, df = Inf, vcov = vcovHC(olsb1r, type=
    "HC")), coeftest(olsb1Femw, df = Inf, vcov = vcovHC(olsb1Femw,
    type="HC")), coeftest(olsb1Malq, df = Inf, vcov = vcovHC(
    olsb1Malq, type="HC"))))

summary(olsb1r)
summary(olsb1Femw)
summary(olsb1Malq)

vif(olsb1r)
vif(olsb1Femw)
vif(olsb1Malq)
```

JS.R

```r
library("Hmisc")
library("MASS")
library("foreign")
library("stargazer")
library("ordinal")
library("VGAM")
library("car")
library("zoo")
library("lmtest")
library("sandwich")
library("AER")
library("systemfit")
library("ivlewbel")
library("mvtnorm")
library("graphics")
library("lmtest")
library("Matrix")
library("robustbase")
library("ivpack")
library("rms")
library("sciplot")

#2. Models including "satisfaction with financial situation" variable
    (proxy)
js_data2 <- happy_data

#Job satisfaction
table(js_data2$st5job)
#Remove missing values and na's
js_data2 <- subset(js_data2, js_data2$st5job != 6 & js_data2$st5job !
    = 9 & js_data2$st5job != 8)

#Age
table(js_data2$ageb)
```

```r
js_data2$ageb[js_data2$ageb >= 70 ] <- 70
js_data2 <- subset(js_data2, js_data2$ageb != 70)
js_data2$agebSquare <- js_data2$ageb ^ 2

#Satisfaction with financial situation
table(js_data2$st5ecny)
#Removing na's from st5ecny
js_data2 <- subset(js_data2, js_data2$st5ecny != 9)
#Changing order of variables to make the interpretation easier
js_data2$st5ecny[js_data2$st5ecny == 5] <- 0
js_data2$st5ecny[js_data2$st5ecny == 1] <- 5
js_data2$st5ecny[js_data2$st5ecny == 4] <- 1
js_data2$st5ecny[js_data2$st5ecny == 2] <- 4
js_data2$st5ecny[js_data2$st5ecny == 3] <- 2
js_data2$st5ecny[js_data2$st5ecny == 4] <- 3
js_data2$st5ecny[js_data2$st5ecny == 5] <- 4
table(js_data2$st5ecny)

#Gender
table(js_data2$sexa)
js_data2$sexa[js_data2$sexa == 2] <- 0
#Coding as dummy vaiable
js_data2$male <- js_data2$sexa
table(js_data2$male)

#Education
table(js_data2$xxlstsch)
#Removing na's, "higher school or vocational school in the old system," "vocational school/commerce school in the old system" due to lack of observation and camparability
js_data2<- subset(js_data2, js_data2$xxlstsch != 99 & js_data2$xxlstsch != 6 & js_data2$xxlstsch != 5 & js_data2$xxlstsch != 4 & js_data2$xxlstsch != 14)
#Combining "ordinary elementary school in the old system", "higher elementary school in the old system", "junior high school/girls' high school in the old system", "junior high school" and "high school" to "high school and below" variable
js_data2$xxlstsch [js_data2$xxlstsch == 1 | js_data2$xxlstsch == 2 | js_data2$xxlstsch == 3 | js_data2$xxlstsch == 8 | js_data2$xxlstsch == 9  ] <- 0
#Combining "university/graduate school in the old system", "university" and "graduate school" to independant variable "college and higher"
js_data2$xxlstsch [js_data2$xxlstsch == 7 |js_data2$xxlstsch == 12 | js_data2$xxlstsch == 13 | js_data2$xxlstsch == 10 | js_data2$xxlstsch == 11 ] <- 1
#Code dummy graduation variable
js_data2$dolstsch [js_data2$dolstsch == 2 | js_data2$dolstsch == 3 | js_data2$dolstsch == 9] <- 0
js_data2$graduatedCollege <- js_data2$dolstsch * js_data2$xxlstsch
table(js_data2$graduatedCollege)

#Marital Status
```

```r
table(js_data2$domarry)
#Removing, na's as well as cohabiting and seperated people from the sample due lack of enough observations
js_data2 <- subset(js_data2, js_data2$domarry != 5 & js_data2$domarry != 6 & js_data2$domarry != 9)
#Create dummy variables
js_data2$never_married <- js_data2$domarry
table(js_data2$never_married)
js_data2$never_married [js_data2$never_married == 1 | js_data2$never_married == 2 ] <- 3
js_data2$divorced <- js_data2$domarry
js_data2$divorced [js_data2$divorced == 1 | js_data2$divorced == 3 | js_data2$divorced == 4] <- 1
table(js_data1$divorced)
js_data2$widowed <- js_data2$domarry
js_data2$widowed [js_data2$widowed == 1 | js_data2$widowed == 4 ] <- 2
table(js_data2$widowed)

#Number of children
table(js_data2$ccnumttl)
#Removing na's
js_data2<- subset(js_data2, js_data2$ccnumttl != 999)
#Combinining 4,5,6,7,8 to category "above 4 children" due to small amount of observations
js_data2$ccnumttl [js_data2$ccnumttl == 5 | js_data2$ccnumttl == 6 | js_data2$ccnumttl == 7 | js_data2$ccnumttl == 8] <- 4
table(js_data2$ccnumttl)

#Commuting
#Variable doct: yes, no, depends
table(js_data2$docmt)
#Erease Na's and "depends" answers
js_data2<- subset(js_data2, js_data2$docmt != 9 & js_data2$docmt != 3)
#Commuting time
table(js_data2$szcmttl)
#Creating new variable by combining both commuting variables. "Identical living and working place" answers were set to 0.
js_data2$szcmttl[js_data2$docmt == 2] <- 0
#Remove Na's and outliers
js_data2 <- subset(js_data2, js_data2$szcmttl != 9999)
#Combine 0-9 minutes
js_data2$szcmttl[js_data2$szcmttl == 1 | js_data2$szcmttl == 2 | js_data2$szcmttl == 3 | js_data2$szcmttl == 4 | js_data2$szcmttl == 5 | js_data2$szcmttl == 6 | js_data2$szcmttl == 7 | js_data2$szcmttl == 8 | js_data2$szcmttl == 9 ] <- 0
#Combine 10-19 minutes
js_data2$szcmttl[js_data2$szcmttl == 10 | js_data2$szcmttl == 11 | js_data2$szcmttl == 12 | js_data2$szcmttl == 13 | js_data2$szcmttl == 14 | js_data2$szcmttl == 15 | js_data2$szcmttl == 16 | js_data2$szcmttl == 17 | js_data2$szcmttl == 18 | js_data2$szcmttl == 19] <- 1
```

```r
#Combine 20-29 minutes
js_data2$szcmttl[js_data2$szcmttl == 20 | js_data2$szcmttl == 21 | js_data2$szcmttl == 22 | js_data2$szcmttl == 23 | js_data2$szcmttl == 24 | js_data2$szcmttl == 25 | js_data2$szcmttl == 26 | js_data2$szcmttl == 27 | js_data2$szcmttl == 28 | js_data2$szcmttl == 29] <- 2
#Combine 30-39 minutes
js_data2$szcmttl[js_data2$szcmttl == 30 | js_data2$szcmttl == 31 | js_data2$szcmttl == 32 | js_data2$szcmttl == 33 | js_data2$szcmttl == 34 | js_data2$szcmttl == 35 | js_data2$szcmttl == 36 | js_data2$szcmttl == 37 | js_data2$szcmttl == 38 | js_data2$szcmttl == 39] <- 3
#Combine 40-49 minutes
js_data2$szcmttl[js_data2$szcmttl == 40 | js_data2$szcmttl == 41 | js_data2$szcmttl == 42 | js_data2$szcmttl == 43 | js_data2$szcmttl == 44 | js_data2$szcmttl == 45 | js_data2$szcmttl == 46 | js_data2$szcmttl == 47 | js_data2$szcmttl == 48 | js_data2$szcmttl == 49] <- 4
#Combine 50-59 minutes
js_data2$szcmttl[js_data2$szcmttl == 50 | js_data2$szcmttl == 51 | js_data2$szcmttl == 52 | js_data2$szcmttl == 53 | js_data2$szcmttl == 54 | js_data2$szcmttl == 55 | js_data2$szcmttl == 56 | js_data2$szcmttl == 57 | js_data2$szcmttl == 58 | js_data2$szcmttl == 59] <- 5
#Combine 60-59 minutes
js_data2$szcmttl[js_data2$szcmttl == 60 | js_data2$szcmttl == 61 | js_data2$szcmttl == 62 | js_data2$szcmttl == 63 | js_data2$szcmttl == 64 | js_data2$szcmttl == 65 | js_data2$szcmttl == 66 | js_data2$szcmttl == 67 | js_data2$szcmttl == 68 | js_data2$szcmttl == 69] <- 6
#Combine 70-79 minutes
js_data2$szcmttl[js_data2$szcmttl == 70 | js_data2$szcmttl == 71 | js_data2$szcmttl == 72 | js_data2$szcmttl == 73 | js_data2$szcmttl == 74 | js_data2$szcmttl == 75 | js_data2$szcmttl == 76 | js_data2$szcmttl == 77 | js_data2$szcmttl == 78 | js_data2$szcmttl == 79] <- 7
#Combine 80-89 minutes
js_data2$szcmttl[js_data2$szcmttl == 80 | js_data2$szcmttl == 81 | js_data2$szcmttl == 82 | js_data2$szcmttl == 83 | js_data2$szcmttl == 84 | js_data2$szcmttl == 85 | js_data2$szcmttl == 86 | js_data2$szcmttl == 87 | js_data2$szcmttl == 88 | js_data2$szcmttl == 89] <- 8
#Combine remaining values to =<90 minutes
js_data2$szcmttl[js_data2$szcmttl >= 90] <- 9
table(js_data2$szcmttl)

#Work time
js_data2 <- subset(js_data2, js_data2$xjobhwk != 888 & js_data2$xjobhwk != 999)
table(js_data2$xjobhwk)
hist(js_data2$xjobhwk)
js_data2$xjobhwkLog <- log(js_data2$xjobhwk)
hist(js_data2$xjobhwkLog)
```

```r

#Type of employement
table(js_data2$tpjob)
#Removing na's and missing values
js_data2<- subset(js_data2, js_data2$tpjob != 9 & js_data2$tpjob !=
    7)
#Create dummy variables
js_data2$executive <- js_data2$tpjob
js_data2$executive [ js_data2$executive >= 2 ] <- 0
table(js_data2$executive)
js_data2$temporary <- js_data2$tpjob
js_data2$temporary [ js_data2$temporary >= 4 | js_data2$temporary <=
    2 ] <- 2
table(js_data2$temporary)
js_data2$dispatched <- js_data2$tpjob
table(js_data2$dispatched)
js_data2$dispatched [ js_data2$dispatched >= 5 | js_data2$dispatched
    <= 3 ] <- 3
js_data2$self_employed <- js_data2$tpjob
table(js_data2$self_employed)
js_data2$self_employed [ js_data2$self_employed >= 6 | js_data2$self_
    employed <= 4 ] <- 4
js_data2$family_worker  <- js_data2$tpjob
js_data2$family_worker [ js_data2$family_worker <= 5] <- 5
table(js_data2$family_worker)

#Goverment agency employee
js_data2$GovermentAgency <- js_data2$szttlsta
table(js_data2$GovermentAgency)
#Moving goverment agency employees to highest category (100)
js_data2$GovermentAgency[js_data2$GovermentAgency == 12] <- 100
#Everyhing else to 0
js_data2$GovermentAgency[js_data2$GovermentAgency <= 99] <- 0
#Moving goverment agency employees to 1
js_data2$GovermentAgency[js_data2$GovermentAgency == 100] <- 1
#Marking GovernmentAgency as categorical variable
js_data2$GovermentAgencyFactor <- factor(js_data2$GovermentAgency,
    labels=c("Normal","GovernmentAgency"))
table(js_data2$GovermentAgency)

#Health (satisfaction with health and physical condition)
table(js_data2$st5hlthy)
#Removing na's
js_data2 <- subset(js_data2, js_data2$st5hlthy != 9)
#Reversing order for easier understanding
js_data2$st5hlthy[js_data2$st5hlthy == 5] <- 0
js_data2$st5hlthy[js_data2$st5hlthy == 1] <- 5
js_data2$st5hlthy[js_data2$st5hlthy == 4] <- 1
js_data2$st5hlthy[js_data2$st5hlthy == 2] <- 4
js_data2$st5hlthy[js_data2$st5hlthy == 3] <- 2
js_data2$st5hlthy[js_data2$st5hlthy == 4] <- 3
js_data2$st5hlthy[js_data2$st5hlthy == 5] <- 4
table(js_data2$st5hlthy)
```

```
#Satisfaction with living place
table(js_data2$st5areay)
#Removing na's
js_data2 <- subset(js_data2, js_data2$st5areay != 9)
#Reversing order for easier understanding
js_data2$st5areay[js_data2$st5areay == 5] <- 0
js_data2$st5areay[js_data2$st5areay == 1] <- 5
js_data2$st5areay[js_data2$st5areay == 4] <- 1
js_data2$st5areay[js_data2$st5areay == 2] <- 4
js_data2$st5areay[js_data2$st5areay == 3] <- 2
js_data2$st5areay[js_data2$st5areay == 4] <- 3
js_data2$st5areay[js_data2$st5areay == 5] <- 4
table(js_data2$st5areay)

#Leisure satisfaction
table(js_data2$st5leisy)
#Removing na's
js_data2 <- subset(js_data2, js_data2$st5leisy != 9)
#Reversing order for easier understanding
js_data2$st5leisy[js_data2$st5leisy == 5] <- 0
js_data2$st5leisy[js_data2$st5leisy == 1] <- 5
js_data2$st5leisy[js_data2$st5leisy == 4] <- 1
js_data2$st5leisy[js_data2$st5leisy == 2] <- 4
js_data2$st5leisy[js_data2$st5leisy == 3] <- 2
js_data2$st5leisy[js_data2$st5leisy == 4] <- 3
js_data2$st5leisy[js_data2$st5leisy == 5] <- 4
table(js_data2$st5leisy)

#Inclusion of different leisure activities
#Television
table(js_data2$hrtv)
#Removing na's
js_data2 <- subset(js_data2, js_data2$hrtv != 999)
#Combining everything above 7 in one category
js_data2$hrtv[js_data2$hrtv >= 7] <- 7

#Number of books (comics,magazines) per month
table(js_data2$fq5read)
#Removing na's
js_data2 <- subset(js_data2, js_data2$fq5read != 9)

#Playing sports
table(js_data2$fqsport)
#Removing na's
js_data2 <- subset(js_data2, js_data2$fqsport != 9)
#Changing order for easier understanding
js_data2$fqsport[js_data2$fqsport == 5] <- 0
js_data2$fqsport[js_data2$fqsport == 1] <- 5
js_data2$fqsport[js_data2$fqsport == 4] <- 1
js_data2$fqsport[js_data2$fqsport == 2] <- 4
js_data2$fqsport[js_data2$fqsport == 3] <- 2
js_data2$fqsport[js_data2$fqsport == 4] <- 3
```

```
js_data2$fqsport[js_data2$fqsport == 5] <- 4
table(js_data2$fqsport)

#Meeting friends (frequency)
table(js_data2$fq7frsee)
#Removing na's
js_data2 <- subset(js_data2, js_data2$fq7frsee != 9)
#Combining 1(every day), 2(several times a week) to "several times a
    week"
js_data2$fq7frsee[js_data2$fq7frsee == 1 ] <- 2
#Changing order for better undrstanding
js_data2$fq7frsee[js_data2$fq7frsee == 7] <- 0
js_data2$fq7frsee[js_data2$fq7frsee == 6] <- 1
js_data2$fq7frsee[js_data2$fq7frsee == 2] <- 6
js_data2$fq7frsee[js_data2$fq7frsee == 5] <- 2
js_data2$fq7frsee[js_data2$fq7frsee == 3] <- 5
js_data2$fq7frsee[js_data2$fq7frsee == 4] <- 3
js_data2$fq7frsee[js_data2$fq7frsee == 5] <- 4
js_data2$fq7frsee[js_data2$fq7frsee == 6] <- 5
table(js_data2$fq7frsee)

#Civic group member
#Combining membership in citizens movement/consumers
    cooperative groups and membership in social service group
table(js_data2$memcivil)
js_data2$civic <- js_data2$memcivil
table(js_data2$civic)
js_data2$civic[js_data2$civic == 2] <- 0
#Combining membership and volunteer organisation and membership in
    civil society
js_data2$civic[js_data2$memvlntr == 1] <- 1
#Removing na's
js_data2 <- subset(js_data2, js_data2$civic != 9)
table(js_data2$civic)

#Hobby group member
table(js_data2$memhobby)
js_data2 <- subset(js_data2, js_data2$memhobby != 9)
js_data2$memhobby[js_data2$memhobby == 2] <- 0
table(js_data2$memhobby)

#Barplot for job satifaction distribution
barplot(table(js_data2$st5job), col="black", names.arg = c("satisfied
     (1)", "somewhat satisfied (2)", "neither (3)","somewhat
    dissatisfied (4)", "dissatisfied (5)"), cex.axis=0.8, cex.names
    =0.7, xlab= "Level of job dissatisfaction", ylab="Number of
    responses", main=NULL)

#Combining last categories of job satisfaction and changing order for
     better understanding
js_data2$st5job[js_data2$st5job == 5 | js_data2$st5job == 4 ] <- 0
#Changing order for better understanding
js_data2$st5job[js_data2$st5job == 1 ] <- 4
```

```
js_data2$st5job[js_data2$st5job == 3 ] <- 1
js_data2$st5job[js_data2$st5job == 4 ] <- 3
table(js_data2$st5job)

#Barplot for new distribution of job satifaction
barplot(table(js_data2$st5job), col="black", names.arg = c( "(somewhat) dissatisfied (0)", "neither (1)", "somewhat satisfied (2)", "satisfied (3)"), cex.axis=0.8, cex.names=0.8, xlab= "Level of job satisfaction", ylab="Number of responses", main=NULL)

#Descriptive Statistics
summary(js_data2$st5job)
se(js_data2$st5job)
summary(js_data2$ageb)
se(js_data2$ageb)
summary(js_data2$male)
se(js_data2$male)
summary(js_data2$never_married)
se(js_data2$never_married)
summary(js_data2$divorced)
se(js_data2$divorced)
summary(js_data2$widowed)
se(js_data2$widowed)
summary(js_data2$ccnumttl)
se(js_data2$ccnumttl)
summary(js_data2$graduatedCollege)
se(js_data2$graduatedCollege)
summary(js_data2$st5ecny)
se(js_data2$st5ecny)
summary(js_data2$xjobhwk)
se(js_data2$xjobhwk)
summary(js_data2$executive)
se(js_data2$executive)
summary(js_data2$temporary)
se(js_data2$temporary)
summary(js_data2$dispatched)
se(js_data2$dispatched)
summary(js_data2$self_employed)
se(js_data2$self_employed)
summary(js_data2$family_worker)
se(js_data2$family_worker)
summary(js_data2$GovermentAgency)
se(js_data2$GovermentAgency)
summary(js_data2$szcmttl)
se(js_data2$szcmttl)
summary(js_data2$st5hlthy)
se(js_data2$st5hlthy)
summary(js_data2$st5areay)
se(js_data2$st5areay)
summary(js_data2$st5leisy)
se(js_data2$st5leisy)
summary(js_data2$civic)
se(js_data2$civic)
```

```
summary(js_data2$memhobby)
se(js_data2$memhobby)
summary(js_data2$fq7frsee)
se(js_data2$fq7frsee)
summary(js_data2$fqsport)
se(js_data2$fqsport)
summary(js_data2$fq5read)
se(js_data2$fq5read)
summary(js_data2$hrtv)
se(js_data2$hrtv)


#Extimating regression for main sample (proxy models)
ols2 <- lm(st5job ~ ageb + agebSquare + male + never_married +
    divorced + widowed + ccnumttl + graduatedCollege + st5ecny +
    xjobhwkLog + executive + temporary + dispatched + self_employed +
     family_worker + GovermentAgency + szcmttl + st5hlthy + st5areay+
     st5leisy + civic + memhobby + fq7frsee + fqsport + fq5read +
    hrtv, data=js_data2)
summary(ols2)

#Breusch-Pegan test
bptest(ols2)
#Result: BP = 65.2198, df = 26, p-value = 3.208e-05
#Estimation of model with heteroskedasticity robust standard errors
coeftest(ols2, df = Inf, vcov = vcovHC(ols2, type="HC"))

#Test for multicollianrity
vif(ols2)

#Correlation Matrix
attach(js_data2)
d1 <- data.frame(ageb, agebSquare, male, never_married, divorced,
    widowed, ccnumttl, graduatedCollege, szincomx, xjobhwkLog,
    executive, temporary, dispatched, self_employed, family_worker,
    GovermentAgency, szcmttl, st5hlthy, st5areay, st5leisy, civic,
    memhobby, fq7frsee, fqsport, fq5read, hrtv)
cor(d1) # get correlations
as.matrix(cor(d1))
detach(js_data2)

#Estimating Model acounding for hetereskedasticity using sandwitch
    estimator
(ordered_logit2 <- lrm(st5job ~ ageb + agebSquare + male + never_
    married + divorced + widowed + ccnumttl + graduatedCollege +
    st5ecny + xjobhwkLog + executive + temporary + dispatched + self_
    employed + family_worker + GovermentAgency + szcmttl + st5hlthy +
     st5areay+ st5leisy + civic + memhobby + fq7frsee + fqsport +
    fq5read + hrtv, x=TRUE, y=TRUE, data=js_data2))
bptest(ordered_logit2)
#Result: BP = 64.7719, df = 25, p-value = 2.23e-05
(g2 <- robcov(ordered_logit2))
```

```
#Calculate McFadden Pseudo-R2
vglm11 <- vglm(st5job ~ 1, family=propodds, data=js_data2)
LLf   <- logLik(g2)
LL0   <- logLik(vglm11)
as.vector(1 - (LLf / LL0))

#Calculate Loglikelihood
(logLik(g2))

#VIFs for multicolianarity
vif(g2)


#Estimation of regression for main famale sample
js_data2Fem <- js_data2
#Ereasing male sample
js_data2Fem <- subset(js_data2Fem,js_data2Fem$sexa != 1)
#Estimation OLS
ols2Fem <- lm(st5job ~ ageb + agebSquare + never_married + divorced +
     widowed + ccnumttl +  graduatedCollege + st5ecny + xjobhwkLog +
    executive + temporary + dispatched + self_employed + family_
    worker + GovermentAgency + szcmttl + st5hlthy + st5areay+
    st5leisy + civic + memhobby + fq7frsee + fqsport + fq5read + hrtv
     , data=js_data2Fem)
summary(ols2Fem)

#Descriptive Statistics
summary(js_data2Fem$st5job)
se(js_data2Fem$st5job)
summary(js_data2Fem$ageb)
se(js_data2Fem$ageb)
summary(js_data2Fem$never_married)
se(js_data2Fem$never_married)
summary(js_data2Fem$divorced)
se(js_data2Fem$divorced)
summary(js_data2Fem$widowed)
se(js_data2Fem$widowed)
summary(js_data2Fem$ccnumttl)
se(js_data2Fem$ccnumttl)
summary(js_data2Fem$graduatedCollege)
se(js_data2Fem$graduatedCollege)
summary(js_data2Fem$st5ecny)
se(js_data2Fem$st5ecny)
summary(js_data2Fem$xjobhwk)
se(js_data2Fem$xjobhwk)
summary(js_data2Fem$executive)
se(js_data2Fem$executive)
summary(js_data2Fem$temporary)
se(js_data2Fem$temporary)
summary(js_data2Fem$dispatched)
se(js_data2Fem$dispatched)
summary(js_data2Fem$self_employed)
se(js_data2Fem$self_employed)
```

```
summary(js_data2Fem$family_worker)
se(js_data2Fem$family_worker)
summary(js_data2Fem$GovermentAgency)
se(js_data2Fem$GovermentAgency)
summary(js_data2Fem$szcmttl)
se(js_data2Fem$szcmttl)
summary(js_data2Fem$st5hlthy)
se(js_data2Fem$st5hlthy)
summary(js_data2Fem$st5areay)
se(js_data2Fem$st5areay)
summary(js_data2Fem$st5leisy)
se(js_data2Fem$st5leisy)
summary(js_data2Fem$civic)
se(js_data2Fem$civic)
summary(js_data2Fem$memhobby)
se(js_data2Fem$memhobby)
summary(js_data2Fem$fq7frsee)
se(js_data2Fem$fq7frsee)
summary(js_data2Fem$fqsport)
se(js_data2Fem$fqsport)
summary(js_data2Fem$fq5read)
se(js_data2Fem$fq5read)
summary(js_data2Fem$hrtv)
se(js_data2Fem$hrtv)

#Breusch-Pagan test
bptest(ols2Fem)
#Results: BP = 54.7098, df = 25, p-value = 0.0005371
#Estimation of model with heteroskedasticity robust standard errors
coeftest(ols2Fem, df = Inf, vcov = vcovHC(ols2Fem, type="HC"))

#Test for multicollianrity
vif(ols2Fem)

#Correlation Matrix
attach(js_data2Fem)
ols2Femcor <- data.frame(ageb, agebSquare, never_married, divorced,
    widowed, ccnumttl, graduatedCollege, st5ecny, xjobhwkLog,
    executive, temporary, dispatched, self_employed, family_worker,
    GovermentAgency, szcmttl, st5hlthy, st5areay, st5leisy, civic,
    memhobby, fq7frsee, fqsport, fq5read, hrtv)
cor(ols2Femcor) # get correlations
as.matrix(cor(ols2Femcor))
detach(js_data2Fem)


#Extimation ordered logit female sample
(ordered_logit2Fem <- lrm(st5job ~ ageb + agebSquare + never_married
    + divorced + widowed + ccnumttl + graduatedCollege + st5ecny +
    xjobhwkLog + executive + temporary + dispatched + self_employed +
     family_worker + GovermentAgency + szcmttl + st5hlthy + st5areay+
     st5leisy + civic + memhobby + fq7frsee + fqsport + fq5read +
    hrtv, x=TRUE, y=TRUE, data=js_data2Fem))
```

```r

#Estimating Model acounding for hetereskedasticity using sandwitch estimator
bptest(ordered_logit2Fem)
#Results:BP = 57.3278, df = 24, p-value = 0.0001507
(f2 <- robcov(ordered_logit2Fem))

#Calculate McFadden Pseudo-R2
vglm22 <- vglm(st5job ~ 1, family=propodds, data=js_data2Fem)
LLf   <- logLik(f2)
LL0   <- logLik(vglm22)
as.vector(1 - (LLf / LL0))

#Calculate Loglikelihood
(logLik(f2))

#VIFs for multicolianarity check
vif(f2)


#Estmation of regression for male sample
js_data2Mal <- js_data2
#Ereasing female sample
js_data2Mal <- subset(js_data2Mal, js_data2Mal$sexa != 0)
#Estimation OLS
ols2Mal <- lm(st5job ~ ageb + agebSquare + never_married + divorced +
    widowed + ccnumttl + graduatedCollege + st5ecny + xjobhwkLog +
    executive + temporary + dispatched + self_employed + family_
    worker + GovermentAgency + szcmttl + st5hlthy + st5areay +
    st5leisy + civic + memhobby + fq7frsee + fqsport + fq5read + hrtv
    , data=js_data2Mal)
summary(ols2Mal)

#Descriptive Statistics
summary(js_data2Mal$st5job)
se(js_data2Mal$st5job)
summary(js_data2Mal$ageb)
se(js_data2Mal$ageb)
summary(js_data2Mal$male)
se(js_data2Mal$male)
summary(js_data2Mal$never_married)
se(js_data2Mal$never_married)
summary(js_data2Mal$divorced)
se(js_data2Mal$divorced)
summary(js_data2Mal$widowed)
se(js_data2Mal$widowed)
summary(js_data2Mal$ccnumttl)
se(js_data2Mal$ccnumttl)
summary(js_data2Mal$graduatedCollege)
se(js_data2Mal$graduatedCollege)
summary(js_data2Mal$st5ecny)
se(js_data2Mal$st5ecny)
summary(js_data2Mal$xjobhwk)
```

```
se(js_data2Mal$xjobhwk)
summary(js_data2Mal$executive)
se(js_data2Mal$executive)
summary(js_data2Mal$temporary)
se(js_data2Mal$temporary)
summary(js_data2Mal$dispatched)
se(js_data2Mal$dispatched)
summary(js_data2Mal$self_employed)
se(js_data2Mal$self_employed)
summary(js_data2Mal$family_worker)
se(js_data2Mal$family_worker)
summary(js_data2Mal$GovermentAgency)
se(js_data2Mal$GovermentAgency)
summary(js_data2Mal$szcmttl)
se(js_data2Mal$szcmttl)
summary(js_data2Mal$st5hlthy)
se(js_data2Mal$st5hlthy)
summary(js_data2Mal$st5areay)
se(js_data2Mal$st5areay)
summary(js_data2Mal$st5leisy)
se(js_data2Mal$st5leisy)
summary(js_data2Mal$civic)
se(js_data2Mal$civic)
summary(js_data2Mal$memhobby)
se(js_data2Mal$memhobby)
summary(js_data2Mal$fq7frsee)
se(js_data2Mal$fq7frsee)
summary(js_data2Mal$fqsport)
se(js_data2Mal$fqsport)
summary(js_data2Mal$fq5read)
se(js_data2Mal$fq5read)
summary(js_data2Mal$hrtv)
se(js_data2Mal$hrtv)

#Breusch-Pagan test
bptest(ols2Mal)
#Results: BP = 29.5184, df = 25, p-value = 0.2428
#Estimation of model with heteroskedasticity robust standard errors
coeftest(ols2Mal, df = Inf, vcov = vcovHC(ols1Mal, type="HC"))

#Test for multicollianrity
vif(ols2Mal)

#Correlation Matrix
attach(js_data2Mal)
ols2Malcor <- data.frame(ageb, agebSquare, never_married, divorced,
    widowed, ccnumttl, graduatedCollege, st5ecny, xjobhwkLog,
    executive, temporary, dispatched, self_employed, family_worker,
    GovermentAgency, szcmttl, st5hlthy, st5areay, st5leisy, civic,
    memhobby, fq7frsee, fqsport, fq5read, hrtv)
cor(ols2Malcor) #get correlations
as.matrix(cor(ols2Malcor))
detach(js_data2Mal)
```

```
#Estimation ordered logit
(ordered_logit2Mal<- lrm(st5job ~ ageb + agebSquare + never_married + divorced + widowed + ccnumttl + graduatedCollege + st5ecny + xjobhwkLog + executive + temporary + dispatched + self_employed + family_worker + GovermentAgency + szcmttl + st5hlthy + st5areay+ st5leisy + civic + memhobby + fq7frsee + fqsport + fq5read + hrtv,x=TRUE, y=TRUE, data=js_data2Mal))

#Estimating Model acounding for hetereskedasticity using sandwitch estimator
bptest(ordered_logit2Mal)
#Results:BP = 28.066, df = 24, p-value = 0.2573
(m2 <- robcov(ordered_logit2Mal))

#VIFs for multicolianarity check
vif(m2)

#Calculate McFadden Pseudo-R2
vglm33 <- vglm(st5job ~ 1, family=propodds, data=js_data2Mal)
LLf   <- logLik(m2)
LL0   <- logLik(vglm33)
as.vector(1 - (LLf / LL0))

#Calculate Loglikelihood
(logLik(m2))

#Request OLS Main (Proxy Models)
stargazer(list(coeftest(ols2, df = Inf, vcov = vcovHC(ols2, type="HC")), coeftest(ols2Fem, df = Inf, vcov = vcovHC(ols2Fem, type="HC")),coeftest(ols2Mal, df = Inf, vcov = vcovHC(ols2Mal, type="HC"))))
#Request Ordered Logit Main
stargazer(list(g2, f2, m2))

(ordered_logit2 <- lrm(st5job ~ ageb + agebSquare + st5ecny + sexaFactor + xxlstschFactor + domarryFactor + ccnumttl + szcmttl + xjobhwkLog + tpjobFactor +  GovermentAgencyFactor + st5hlthy + st5areay + st5leisy + civic + fq7frsee + fqsport + fq5read + hrtv, data=js_data2))

lewbel(st5job ~  st5leisy  + st5ecny + st5areay + st5hlthy  + civic + fq7frsee + fqsport + fq5read + hrtv + xxlstsch + domarry| ageb + agebSquare  + sexaFactor  + szcmttl  + ccnumttl + xjobhwkLog + tpjobFactor + GovermentAgencyFactor | ageb + agebSquare + sexaFactor  + szcmttl  + ccnumttl + xjobhwkLog + tpjobFactor + GovermentAgencyFactor , data=js_data2, clustervar = NULL, robust = TRUE)

#Estimation of regression for main famale sample
js_data2Fem <- js_data2
#Ereasing male sample
```

```
js_data2Fem <- subset(js_data2Fem,js_data2Fem$sexa != 1)
#Estimation OLS
ols2Fem <- lm(st5job ~ ageb + agebSquare + st5ecny + xxlstschFactor +
     domarryFactor + ccnumttl + szcmttl + xjobhwkLog + tpjobFactor +
    GovermentAgencyFactor + st5hlthy + st5areay + st5leisy + civic +
    memhobby + fq7frsee + fqsport + fq5read + hrtv , data=js_data2Fem
    )
summary(ols2Fem)
#Extimation ordered logit
(ordered_logit2Fem <- lrm(st5job ~ ageb + agebSquare + st5ecny +
    xxlstschFactor + domarryFactor + ccnumttl + szcmttl + xjobhwkLog
    + tpjobFactor + GovermentAgencyFactor + st5hlthy + st5areay +
    st5leisy + civic + memhobby + fq7frsee + fqsport + fq5read + hrtv
     , data=js_data2Fem))

#Estmation of regression for male sample
js_data2Mal <- js_data2
#Ereasing female sample
js_data2Mal <- subset(js_data2Mal, js_data2Mal$sexa != 0)
#Estimation OLS
ols2Mal <- lm(st5job ~ ageb + agebSquare + st5ecny + xxlstschFactor +
     domarryFactor + ccnumttl + szcmttl + xjobhwkLog + tpjobFactor +
    GovermentAgencyFactor + st5hlthy + st5areay + st5leisy + civic +
    memhobby + fq7frsee + fqsport + fq5read + hrtv , data=js_data2Mal
    )
summary(ols2Mal)
#Extimation ordered logit
(ordered_logit2Mal<- lrm(st5job ~ ageb + agebSquare + st5ecny +
    xxlstschFactor + domarryFactor + ccnumttl + szcmttl + xjobhwkLog
    + tpjobFactor + GovermentAgencyFactor + st5hlthy + st5areay +
    st5leisy + civic + memhobby + fq7frsee + fqsport + fq5read + hrtv
     , data=js_data2Mal))

#Request output 5
stargazer(list(ols2, ols2Fem, ols2Mal))
#Request output 6
stargazer(list(ordered_logit2, ordered_logit2Fem, ordered_logit2Mal))


#Models derived from Form B
js_data2b <- js_data2

#Leisure activities
#Time in nature (woods, ocean, or river)
table(js_data2b$frlxnatr)
js_data2b <- subset(js_data2b, js_data2b$frlxnatr != 9 &  js_data2b$
    frlxnatr !=8)
#Combining 1,2,3 "to at least once in weak" category
js_data2b$frlxnatr[js_data2b$frlxnatr == 2 | js_data2b$frlxnatr == 1]
     <- 3
#Changing order for better understanding
js_data2b$frlxnatr[js_data2b$frlxnatr == 6] <- 0
js_data2b$frlxnatr[js_data2b$frlxnatr == 5] <- 1
```

```
js_data2b$frlxnatr[js_data2b$frlxnatr == 4] <- 2
table(js_data2b$frlxnatr)

#Listening to music or sing a song
table(js_data2b$frlxmsc)
js_data2b <- subset(js_data2b, js_data2b$frlxmsc != 9)
#Changing order for better understanding
js_data2b$frlxmsc[js_data2b$frlxmsc == 6] <- 0
js_data2b$frlxmsc[js_data2b$frlxmsc == 1] <- 6
js_data2b$frlxmsc[js_data2b$frlxmsc == 5] <- 1
js_data2b$frlxmsc[js_data2b$frlxmsc == 6] <- 5
js_data2b$frlxmsc[js_data2b$frlxmsc == 2] <- 6
js_data2b$frlxmsc[js_data2b$frlxmsc == 4] <- 2
js_data2b$frlxmsc[js_data2b$frlxmsc == 6] <- 4
table(js_data2b$frlxmsc)

#Soldarity at work place
table(js_data2b$opwslpw)
#Removing na's
js_data2b <- subset(js_data2b, js_data2b$opwslpw != 9 & js_data2b$
    opwslpw != 5)
#Changing order for better understanding
js_data2b$opwslpw[js_data2b$opwslpw == 4] <- 0
js_data2b$opwslpw[js_data2b$opwslpw == 1] <- 4
js_data2b$opwslpw[js_data2b$opwslpw == 3] <- 1
js_data2b$opwslpw[js_data2b$opwslpw == 4] <- 3


#Extimating Regression for sub-sample
ols2b <- lm(st5job ~ ageb + agebSquare + st5ecny + sexaFactor +
    xxlstschFactor + domarryFactor + ccnumttl + szcmttl + xjobhwkLog
    + opwslpw + tpjobFactor + GovermentAgencyFactor + st5hlthy +
    st5areay + st5leisy  + frlxmsc  + frlxnatr + civic  + memhobby +
    fq7frsee + fqsport + fq5read + hrtv , data=js_data2b)
summary(ols2b)
(ordered_logit2b <- lrm(st5job ~ ageb + agebSquare + st5ecny +
    sexaFactor + xxlstschFactor + domarryFactor + ccnumttl + szcmttl
    + xjobhwkLog + opwslpw + tpjobFactor + GovermentAgencyFactor +
    st5hlthy + st5areay + st5leisy  + frlxmsc  + frlxnatr + civic +
    memhobby + fq7frsee + fqsport + fq5read + hrtv , data=js_data2b))

#Estimation of regression for famale sample
js_data2bFem <- js_data2b
#Ereasing male sample
js_data2bFem <- subset(js_data2bFem,js_data2bFem$sexa != 1)
#Estimation OLS
ols2bFem <- lm(st5job ~ ageb + agebSquare + st5ecny + xxlstschFactor
    + domarryFactor + ccnumttl + szcmttl + xjobhwkLog + opwslpw +
    tpjobFactor + GovermentAgencyFactor + st5hlthy + st5areay +
    st5leisy  + frlxmsc  + frlxnatr + civic + memhobby + fq7frsee +
    fqsport + fq5read + hrtv , data=js_data2bFem)
summary(ols2bFem)
#Extimation ordered logit
```

```
(ordered_logit2bFem <- lrm(st5job ~ ageb + agebSquare + st5ecny +
    xxlstschFactor + domarryFactor + ccnumttl + szcmttl + xjobhwkLog
    + opwslpw + tpjobFactor + GovermentAgencyFactor + st5hlthy +
    st5areay + st5leisy  + frlxmsc  + frlxnatr + civic + memhobby +
    fq7frsee + fqsport + fq5read + hrtv  , data=js_data2bFem))

#Estmation of regression for male sample
js_data2bMal <- js_data2b
#Ereasing female sample
js_data2bMal <- subset(js_data2bMal, js_data2bMal$sexa != 0)
#Estimation OLS
ols2bMal <- lm(st5job ~ ageb + agebSquare + st5ecny + xxlstschFactor
    + domarryFactor + ccnumttl + szcmttl + xjobhwkLog + opwslpw +
    tpjobFactor + GovermentAgencyFactor + st5hlthy + st5areay +
    st5leisy  + frlxmsc  + frlxnatr + civic + memhobby + fq7frsee +
    fqsport + fq5read + hrtv  , data=js_data2bMal)
summary(ols2bMal)
#Extimation ordered logit
(ordered_logit2bMal<- lrm(st5job ~ ageb + agebSquare + st5ecny +
    xxlstschFactor + domarryFactor + ccnumttl + szcmttl + xjobhwkLog
    + opwslpw + tpjobFactor + GovermentAgencyFactor + st5hlthy +
    st5areay + st5leisy  + frlxmsc  + frlxnatr + civic + memhobby +
    fq7frsee + fqsport + fq5read + hrtv  , data=js_data2bMal))

#Request output 7
stargazer(list(ols2b, ols2bFem ,ols2bMal))
#Request output 8
stargazer(list(ordered_logit2b, ordered_logit2bFem, ordered_
    logit2bMal))

#Including happiness
js_data2bhap <- js_data2b

#Happiness
table(js_data2bhap$op5happe)
js_data2bhap <- subset(js_data2bhap, js_data2bhap$op5happe != 9)
#Combine two lowest categories due to lack of observations
js_data2bhap$op5happe[js_data2bhap$op5happe == 5] <- 4
#Change order for better understanding
js_data2bhap$op5happe[js_data2bhap$op5happe == 4] <- 0
js_data2bhap$op5happe[js_data2bhap$op5happe == 1] <- 4
js_data2bhap$op5happe[js_data2bhap$op5happe == 3] <- 1
js_data2bhap$op5happe[js_data2bhap$op5happe == 4] <- 3

ols2bhap <- lm(st5job ~ op5happe + ageb + agebSquare +  st5ecny +
    sexaFactor + xxlstschFactor + domarryFactor + ccnumttl + szcmttl
    + xjobhwkLog + opwslpw + tpjobFactor + GovermentAgencyFactor +
    st5hlthy + st5areay + st5leisy  + frlxmsc  + frlxnatr + civic +
    memhobby + fq7frsee + fqsport + fq5read + hrtv  , data=js_
    data2bhap)
summary(ols2bhap)
(ordered_logit2bhapp <- lrm(st5job ~ op5happe + ageb + agebSquare +
    st5ecny + sexaFactor + xxlstschFactor + domarryFactor + ccnumttl
```

```
+ szcmttl + xjobhwkLog + opwslpw + tpjobFactor +
GovermentAgencyFactor + st5hlthy + st5areay + st5leisy  + frlxmsc
 + frlxnatr + civic + memhobby + fq7frsee + fqsport + fq5read +
hrtv , js_data2bhap))

#Estmation of regression for male sample
js_data2bMalhap <- js_data2bhap
#Ereasing female sample
js_data2bMalhap <- subset(js_data2bMalhap, js_data2bMalhap$sexa != 0)
#Estimation OLS
ols2bhapMal <- lm(st5job ~ op5happe + ageb + agebSquare +st5ecny  +
    xxlstschFactor + domarryFactor + ccnumttl + szcmttl + xjobhwkLog
    + opwslpw + tpjobFactor + GovermentAgencyFactor + st5hlthy +
    st5areay + st5leisy  + frlxmsc  + frlxnatr + civic +  memhobby +
    fq7frsee + fqsport + fq5read + hrtv , data=js_data2bMalhap)
summary(ols2bhapMal)
#Extimation ordered logit
(ordered_logit2bMalhap<- lrm(st5job ~ op5happe + ageb + agebSquare +
    st5ecny  + xxlstschFactor + domarryFactor + ccnumttl + szcmttl +
    xjobhwkLog + opwslpw + tpjobFactor + GovermentAgencyFactor +
    st5hlthy + st5areay + st5leisy  + frlxmsc  + frlxnatr + civic +
    memhobby + fq7frsee + fqsport + fq5read + hrtv , js_data2bMalhap)
    )

#Estmation of regression for female sample
js_data2bFemhap <- js_data2bhap
#Ereasing female sample
js_data2bFemhap <- subset(js_data2bFemhap, js_data2bFemhap$sexa != 1)
#Estimation OLS
ols2bhapFem <- lm(st5job ~ op5happe + ageb + agebSquare +  st5ecny +
    xxlstschFactor + domarryFactor + ccnumttl + szcmttl + xjobhwkLog
    + opwslpw + tpjobFactor + GovermentAgencyFactor + st5hlthy +
    st5areay + st5leisy  + frlxmsc  + frlxnatr + civic  +  memhobby +
     fq7frsee + fqsport + fq5read + hrtv , data=js_data2bFemhap)
summary(ols2bhapFem)
#Extimation ordered logit
(ordered_logit2bFemhap<- lrm(st5job ~ op5happe + ageb + agebSquare +
     st5ecny + xxlstschFactor + domarryFactor + ccnumttl + szcmttl +
    xjobhwkLog + opwslpw + tpjobFactor + GovermentAgencyFactor +
    st5hlthy + st5areay + st5leisy  + frlxmsc  + frlxnatr + civic  +
     memhobby + fq7frsee + fqsport + fq5read + hrtv , js_data2bFemhap
    ))
```

JS_Proxy.R

Bibliography

4.8 instrumental variables. (n.d.). Available from http://cameron.econ.ucdavis.edu/e240a/ch04iv.pdf

Allison, P. (2012). When can you safely ignore multicollinearity? Available from http://www.statisticalhorizons.com/multicollinearity

Allison, P. D. (2012). *Logistic regression using SAS: Theory and application* (Second Edition ed.). SAS Institute.

Angrist, J., & Pischke, J. (2009). *Mostly harmless econometrics: An empiricist's companion.* Princeton University Press.

Antonakis, J., Bendahan, S., Jacquart, P., & Lalive, R. (2010, December). On making causal claims: A review and recommendations. *The Leadership Quarterly*, *21*(6), 1086–1120. Available from http://linkinghub.elsevier.com/retrieve/pii/S1048984310001475

Antonakis, J., Bendahan, S., Jacquart, P., & Lalive, R. (2014). Causality and endogeneity: Problems and solutions. In D. Day (Ed.), *The oxford handbook of leadership and organizations* (pp. 93–117).

Argyle, M. (1999). Causes and correlates. In D. Kahnemann, E. Diener, & N. Schwarz (Eds.), *Wellbeing: the foundations of hedonistic psychology* (pp. 353–373). New York: Sage.

Argyle, M. (2013). *The psychology of happiness.* Routledge.

Asadullah, M. N., & Fernández, R. M. (2008). Work life balance practices and the gender gap in job satisfaction in the UK: Evidence from matched employer employee data. *IZA Discussion Papers*(3582), 1–39. Available from http://ftp.iza.org/dp3582.pdf

Balatsky, G., & Diener, E. (1993). Subjective well-being among Russian students. *Social Indicators Research*, *28*(3), 225–243. Available from http://www.jstor.org/stable/27522672

Ball, R., & Chernova, K. (2007, November). Absolute income, relative income, and happiness. *Social Indicators Research*, *88*(3), 497–529. Available from http://www.springerlink.com/index/10.1007/s11205-007-9217-0

Baltagi, B. H. (2011). *Econometrics* (5th ed.). Springer.

Bauer, T. K. (2004). High performance workplace practices and job satisfaction: Evidence from Europe. *IZA Discussion Paper Series*(1265), 1–33. Available from http://ftp.iza.org/dp1265.pdf

Bender, K. A. (2005, February). Job satisfaction and gender segregation. *Oxford Economic Papers*, *57*(3), 479–496. Available from http://oep.oupjournals.org/cgi/doi/10.1093/oep/gpi015

Benz, M., & Frey, B. S. (2008, May). Being independent is a great thing: Subjective evaluations of self-employment and hierarchy. *Economica*, *75*(298), 362–383. Available from http://doi.wiley.com/10.1111/j.1468-0335.2007.00594.x

Bianchi, M. (2012). Financial development, entrepreneurship, and job satisfaction. *Review of Economics and Statistics*, *94*(February), 273–286.

Blanchflower, D. G., & Oswald, A. J. (2004, July). Well-being over time in Britain and the USA. *Journal of Public Economics*, *88*(7-8), 1359–1386. Available from http://linkinghub.elsevier.com/retrieve/pii/S0047272702001688

Bound, J., Jaeger, D. A., & Baker, R. M. (1995). Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak. *Journal of the American Statistical Association*, *90*(430), 443–450.

Bowling, N. a., Eschleman, K. J., & Wang, Q. (2010, December). A meta-analytic examination of the relationship between job satisfaction and subjective well-being. *Journal of Occupational and Organizational Psychology*, *83*(4), 915–934. Available from http://doi.wiley.com/10.1348/096317909X478557

Boyles, C., & Shibata, A. (2009, January). Job satisfaction, work time, and well-being among married women in japan. *Feminist Economics*, *15*(1), 57–84. Available from http://www.tandfonline.com/doi/abs/10.1080/13545700802629378

Chun, S., Lee, Y., Kim, B., & Heo, J. (2012, October). The contribution of leisure participation and leisure satisfaction to stress-related growth. *Leisure Sciences*, *34*(5), 436–449. Available from http://www.tandfonline.com/doi/abs/10.1080/01490400.2012.714704

Clark, A. E. (1997). Job satisfaction and gender: Why are women so happy at work? *Labour Economics*, *4*, 341–372. Available from http://www.sciencedirect.com/science/article/pii/S0927537197000109

Clark, A. E. (1998). What makes a good job? Evidence from OECD countries. In *Csls conference on the state of living standards and the quality of life in canada* (pp. 1–38). Available from http://www.csls.ca/events/oct98/clark.pdf

Clark, A. E. (2001, May). What really matters in a job? Hedonic measurement using quit data. *Labour Economics*, *8*(2), 223–242. Available from http://linkinghub.elsevier.com/retrieve/pii/S0927537101000318

Clark, A. E. (2003). Inequality-aversion and income mobility: A direct test. *DELTA (Ecole normale supérieure)*.

Clark, A. E. (2010). Work, jobs and well-being across the millennium. In E. Diener, D. Kahneman, & J. Helliwell (Eds.), *International differences in well-being* (pp. 436–468). Available from http://books.google.com/books?hl=de&lr=&id=m99aqwLFrGoC&pgis=1

Clark, A. E., Frijters, P., & Shields, M. (2008). Relative income, happiness, and utility: An explanation for the easterlin paradox and other puzzles. *Journal of Economic Literature*, *46*(1), 95–144. Available from http://pubs.aeaweb.org/doi/abs/10.1257/jel.46.1.95

Clark, A. E., Oswald, A., & Warr, P. (1996, March). Is job satisfaction u-shaped in age? *Journal of Occupational and Organizational Psychology*, *69*(1), 57–81. Available from http://doi.wiley.com/10.1111/j.2044-8325.1996.tb00600.x

Craighead, C. W., Ketchen, D. J., Dunn, K. S., & Hult, G. T. M. (2011, August). Addressing common method variance: Guidelines for survey research on information technology, operations, and supply chain management. *IEEE Transactions on Engineering Management*, *58*(3), 578–588. Available from http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5764502

Cramer, J. (2003). *Logit models from economics and other fields*. New York: Cambridge University Press. Available from www.cambridge.org/9780521815888

Crosby, F. J. (1982). *Relative deprivation and working women*. Oxford University Press New York.

Curto, J. D., & Pinto, J. C. (2011, July). The corrected VIF (CVIF). *Journal of Applied Statistics*, *38*(7), 1499–1507. Available from http://www.tandfonline.com/doi/abs/10.1080/02664763.2010.505956

Delmelle, E. C., Haslauer, E., & Prinz, T. (2013, June). Social satisfaction, commuting and neighborhoods. *Journal of Transport Geography*, *30*, 110–116. Available from http://linkinghub.elsevier.com/retrieve/pii/S0966692313000628

DeMaris, A. (2012). A tutorial in logistic regression. *Journal of Marriage and Family*, *57*(4), 956–968. Available from http://www.jstor.org/stable/353415

D'Addio, A. C., Eriksson, T., & Frijters, P. (2007, October). An analysis of the determinants of job satisfaction when individuals' baseline satisfaction levels may differ. *Applied Economics*, *39*(19), 2413–2423. Available from http://www.tandfonline.com/doi/abs/10.1080/00036840600707357

Eagle, B. W., Miles, E. W., & Icenogle, M. L. (1997). Interrole conflicts and the permeability of work and family domains: Are there gender differences? *Journal of Vocational Behavior*, *50*, 168–184.

Easterlin, R. A. (1974). *Does economic growth improve the human lot? some empirical evidence.* Available from http://graphics8.nytimes.com/images/2008/04/16/business/Easterlin1974.pdf

Eckey, H.-F., Kosfeld, R., & Dreger, C. (2011). *Ökonometrie.* Wiesbaden: Gabler Verlag. Available from http://link.springer.com/10.1007/978-3-8349-7051-0

EFILWC. (2007). *European foundation for the improvement of living and working conditions.* Available from http://www.eurofound.europa.eu/ewco/reports/TN0608TR01/TN0608TR01_7.htm

Ekman, P., Davidson, R. J., & Friesen, W. V. (1990, February). The duchenne smile: Emotional expression and brain physiology. *Journal of personality and social psychology*, *58*(2), 342–53. Available from http://www.ncbi.nlm.nih.gov/pubmed/2319446

Engle, R. F. (1984). Wald, likelihood ratio, and lagrange multiplier test in econometrics. In Z. Griliches & M. D. Intriligator (Eds.), *Handbook of econometrics* (2nd ed., Vol. II, pp. 776–825). Amsterdam: Elsevier Science Publishers. Available from http://www.stern.nyu.edu/rengle/LagrangeMultipliersHandbook_of_Econ__II___Engle.pdf

European Commission. (2007). Task force on core social variables - final report. *Eurostat*, 1–127. Available from http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-RA-07-006/EN/KS-RA-07-006-EN.PDF

Fasang, A., Geerdes, S., Schönmann, K., & Siarov, L. (2007). *Job satisfaction and labour market mobility.* Dublin.

Ferree, M. M. (1976). Working-class jobs: Housework and paid work as sources of satisfaction. *Social Problems*, *23*(4), 431–441. Available from http://www.jstor.org/stable/799853.Accessed:

Ferrer-Carbonell, A., & Frijters, P. (2004). How imortant it methology for the estimates of the determinants of happiness? *The Economic Journal*, *114*(1997), 641–659.

Freakonomics Blog. (2014). Why marry? Part 1: Full transcript. Available from http://freakonomics.com/2014/02/13/why-marry-part-1-full-transcript/

Freeman, R. B. (1978). Job satisfaction as an economic variable. *American Economic Review*, *68*(2), 135–141. Available from http://www.jstor.org/stable/1816677

Frey, B. S. (2008). *Happiness - a revolution in economics.* Cambridge, Massachusetts & London, England: MIT Press.

Frey, B. S., & Stutzer, A. (2002a). The economics of happiness. *World Economics*, *3*(1), 25–41.

Frey, B. S., & Stutzer, A. (2002b). *Happiness and economics.* Princeton, New Jersey: Princeton University Press.

Furnham, A. (1981). Personality and activity preference. *British Journal of Social Psychology*, *20*(1), 57–68. Available from http://dx.doi.org/10.1111/j.2044-8309.1981.tb00474.x

Gandelman, N., & Porzecanski, R. (2011, September). Happiness inequality: How much is reasonable? *Social Indicators Research*, *110*(1), 257–269. Available from http://www.springerlink.com/index/10.1007/s11205-011-9929-z

Graaf-Zijl, M. (2011, December). Job satisfaction and contingent employment. *De Economist*, *160*(2), 197–218. Available from http://link.springer.com/10.1007/s10645-011-9180-7

Graham, C., & Felton, A. (2005, November). Inequality and happiness: Insights from Latin America. *The Journal of Economic Inequality*, *4*(1), 107–122. Available from http://www.springerlink.com/index/10.1007/s10888-005-9009-1

Grandey, A. A., & Cropanzano, R. (1999). The conservation of resources model applied to work – family conflict and strain. *Journal of Vocational Behavior*, *54*, 350–370.

Hamermesh, D. S. (1977). Economic aspects of job satisfaction. *Essays in Labor Market Analysis. New York: John Wiley*, 53–72.

Hanson, S. L., & Sloane, M. (1992). Young children and job satisfaction. *Journal of Marriage and Family*, *54*(4), 799–811.

Headey, B., & Wearing, A. (n.d.). Melbourne, Australia.

Hills, P., & Argyle, M. (2001, March). Happiness, introversion–extraversion and happy introverts. *Personality and Individual Differences*, *30*(4), 595–608. Available from http://linkinghub.elsevier.com/retrieve/pii/S0191886900000581

Hodson, R. (1989). Differences in job satisfaction: Why aren't women more dissatisfied? *The Sociological Quarterly*, *30*(3), 385–399.

Hooff, M. L. van, Geurts, S. a.E., Beckers, D. G., & Kompier, M. a.J. (2011, January). Daily recovery from work: The role of activities, effort and pleasure. *Work & Stress*, *25*(1), 55–74. Available from http://www.tandfonline.com/doi/abs/10.1080/02678373.2011.570941

Horley, J., & Lavery, J. J. (1995). Subjective well-being and age. *Social Indicators Research*, *34*(2), 275-282. Available from http://www.jstor.org/stable/27522804

Howard, D. R. (1976). Multivariate relationships between leisure activities and personality. *Research Quarterly. American Alliance for Health, Physical Education and Recreation*, *47*(2), 226-237. Available from http://www.tandfonline.com/doi/abs/10.1080/10671315.1976.10615366

Hu, B., Shao, J., & Palta, M. (2006). Pseudo-R^2 in logistic regression model. *Statistica Sinica*, *16*, 847–860. Available from http://www3.stat.sinica.edu.tw/statistica/oldpdf/a16n39.pdf

Inglehart, R. (1996). The diminishing utility of economic growth: From maximizing security toward maximizing subjective well-being. *Critical review*, *10*(4), 509–531.

Iwasaki, Y. (2001). Contributions of leisure to coping with daily hassles in university students' lives. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, *33*(2), 128–141.

Jones, M. K., Jones, R. J., Latreille, P. L., & Sloane, P. J. (2009). Workplace job satisfaction in Britain: Evidence from linked employer employee data.

IZA Discussion paper series(3677), 1–42. Available from http://ftp.iza.org/dp4101.pdf

Judge, T. a., & Watanabe, S. (1993). Another look at the job satisfaction-life satisfaction relationship. *Journal of Applied Psychology*, *78*(6), 939–948. Available from http://doi.apa.org/getdoi.cfm?doi=10.1037/0021-9010.78.6.939

Kalleberg, A. L., & Vaisey, S. (2005, September). Pathways to a good job: Perceived work quality among the machinists in North America. *British Journal of Industrial Relations*, *43*, 431–454.

Katchova, A. (2013). *Instrumental variables.* Available from https://docs.google.com/file/d/0BwogTI8d6EEiMjRoUG1EdGJic1k/edit

Kawada, T., & Otsuka, T. (2011). Relationship between job stress, occupational position and job satisfaction using a brief job stress questionnaire (BJSQ). *Work (Reading, Mass.)*, *40*(4), 393–399. Available from http://www.ncbi.nlm.nih.gov/pubmed/22130057

Kenny, D. A. (1979). *Correlation and causality.* New York: Wiley-Interscience.

Korpela, K., & Kinnunen, U. (2010, December). How is leisure time interacting with nature related to the need for recovery from work demands? testing multiple mediators. *Leisure Sciences*, *33*(1), 1–14. Available from http://www.tandfonline.com/doi/abs/10.1080/01490400.2011.533103

Lee, T. H., Trevor, C. O., Gerhart, B., & Weller, I. (2008). understanding voluntary turnover: Path-specific job satisfaction effects and the importance of unsolicited job offers. *Academy of Management Journal*, *51*(4), 651–671.

Lewbel, A. (2012, January). Using heteroscedasticity to identify and estimate mismeasured and endogenous regressor models. *Journal of Business & Economic*

Statistics, *30*(1), 67–80. Available from http://www.tandfonline.com/doi/abs/10.1080/07350015.2012.643126

Lincoln, J. R. (1992). *Culture, control, and commitment: A study of work organization and work attitudes in the United States and Japan*. CUP Archive.

Linz, S. J. (2004, July). Motivating Russian workers: analysis of age and gender differences. *The Journal of Socio-Economics*, *33*(3), 261–289. Available from http://linkinghub.elsevier.com/retrieve/pii/S1053535703001240

Linz, S. J., & Semykina, A. (2012, October). What makes workers happy? anticipated rewards and job satisfaction. *Industrial Relations: A Journal of Economy and Society*, *51*(4), 811–844. Available from http://doi.wiley.com/10.1111/j.1468-232X.2012.00702.x

Loscocco, K. A. (1990, May). Reactions to blue-collar work: A comparison of women and men. *Work and Occupations*, *17*(2), 152–177. Available from http://wox.sagepub.com/cgi/doi/10.1177/0730888490017002002

Loscocco, K. A., & Bose, C. E. (1998). Gender and job satisfaction in urban China: The early Post-Mao period. *Social Science Quarterly*, *79*(1), 91–109.

Lounsbury, J. W., Moffitt, L., Gibson, L. W., Drost, A. W., & Stevens, M. (2007, March). An investigation of personality traits in relation to job and career satisfaction of information technology professionals. *Journal of Information Technology*, *22*(2), 174–183. Available from http://www.palgrave-journals.com/doifinder/10.1057/palgrave.jit.2000094

Lu, L., & Hu, C.-H. (2005, September). Personality, leisure experiences and happiness. *Journal of Happiness Studies*, *6*(3), 325–342. Available from http://link.springer.com/10.1007/s10902-005-8628-3

Lu, L., & Kao, S.-F. (2009). Leisure participation and leisure satisfaction: Moderating effects of personality traits. *Journal of Sport and*

Recreation Research, *3*(3), 1–11. Available from http://web.ba.ntu.edu.tw/luolu/Leisure%20Participation%20and%20Leisure%20Satisfaction%20Moderating%20Effects%20of%20Personality%20Traits.pdf

Luttner, E. F. P. (2005, August). Neighbors as negatives: relative earnings and well-being. *The Quarterly Journal of Economics*, 963–1002.

Lynch, S. M. (2003, March). Missing data. , 1–10. Available from http://www.princeton.edu/~slynch/soc504/missingdata.pdf

Mannheim, B. (1983, November). Male and female industrial workers: Job satisfaction, work role centrality, and work place preference. *Work and Occupations*, *10*(4), 413–436. Available from http://wox.sagepub.com/cgi/doi/10.1177/0730888483010004003

Marchenko, Y. V., & Genton, M. G. (2012, March). A Heckman selection-t model. *Journal of the American Statistical Association*, *107*(497), 304–317. Available from http://www.tandfonline.com/doi/abs/10.1080/01621459.2012.656011

Martinez, M. N. (1994). Close link between work and leisure. *HR Magazine*, *39*(8), 25. Available from http://search.ebscohost.com/login.aspx?direct=true&db=buh&AN=9410315768&lang=de&site=ehost-live

Martz, E. (2013). When more is not better - beware of multicollinearity. *Quality*, *52*(8), 24–26.

Mathieu, J. E., & Zajac, D. M. (1990). A review and meta-analysis of the antecedents, correlates, and consequences of organizational commitment. *Psychological Bulletin*, *108*(2), 171–194. Available from http://doi.apa.org/getdoi.cfm?doi=10.1037/0033-2909.108.2.171

McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), *Frontiers in econometrics* (pp. 105–142). Academic Press, New York.

Mobley, W. H. (1977). Intermediate linkages in the relationship between job satisfaction and employee turnover. *Journal of Applied Psychology*, *62*(2), 237–240.

Mojza, E. J., & Sonnentag, S. (2010). Does volunteer work during leisure time buffer negative effects of job stressors? A diary study. *European Journal of Work and Organizational Psychology*, *19*(2), 231–252.

Near, J. P., Rice, R. W., & Hunt, R. G. (1978). Work and extra-work correlates of life and job satisfaction. *The Academy of Management Journal*, *21*(2), 248–264.

Ng, Y.-K. (1997). A case for happiness, cardinalism, and inerpersonal comparability. *The Economic Journal*, *107*, 1848–1858. Available from http://onlinelibrary.wiley.com/doi/10.1111/j.1468-0297.1997.tb00087.x/abstract

Nielsen, I., & Smyth, R. (2008, October). Job satisfaction and response to incentives among China's urban workforce. *The Journal of Socio-Economics*, *37*(5), 1921–1936. Available from http://linkinghub.elsevier.com/retrieve/pii/S105353570800019X

OECD. (2012). *Ageing and employment policies - statistics on average effective age of retirement.* Available from http://www.oecd.org/els/public-pensions/ageingandemploymentpolicies-statisticsonaverageeffectiveageofretirement.htm

OECD. (2013, March). *Guidelines on measuring subjective well-being* (Tech. Rep.). Available from http://www.oecd-ilibrary.org/economics/oecd-guidelines-on-measuring-subjective-well-being_9789264191655-en

Oerlemans, W. G., Bakker, A. B., & Demerouti, E. (2014). How feeling happy during off-job activities helps successful recovery from work: A day reconstruction study. *Work and Stress*, *28*(2), 198-216. Available from http://dx.doi.org/10.1080/02678373.2014.901993

Oshio, T., & Kobayashi, M. (2010, September). Area-level income inequality and individual happiness: Evidence from Japan. *Journal of Happiness Studies*, *12*(4), 633–649. Available from http://www.springerlink.com/index/10.1007/s10902-010-9220-z

Oshio, T., Nozaki, K., & Kobayashi, M. (2010, November). Relative income and happiness in Asia: Evidence from nationwide surveys in China, Japan, and Korea. *Social Indicators Research*, *104*(3), 351–367. Available from http://www.springerlink.com/index/10.1007/s11205-010-9754-9

Pavot, W., & Diener, E. (1993). Review of the satisfaction with life scale. *Psychological Assessment*, *5*(2), 164–172. Available from http://doi.apa.org/getdoi.cfm?doi=10.1037/1040-3590.5.2.164

Pearson, Q. M. (1998). Job satisfaction, leisure satisfaction, and psychological health. *Career Development Quartely*, *46*, 416–427.

Pearson, Q. M. (2008, January). Role overload, job satisfaction, leisure satisfaction, and psychological health among employed women. *Journal of Counseling & Development*, *86*(1), 57–63. Available from http://doi.wiley.com/10.1002/j.1556-6678.2008.tb00626.x

Pfaff, S. (2014). Pendelentfernung, Lebenszufriedenheit und Entlohnung. *Zeitschrift für Soziologie*, *43*(2), 113–130.

Poggi, A. (2014). Within-establishment wage inequality and satisfaction. *Economics: The Open-Access, Open-Assessment E-Journal*, *8*(2014-4). Available from http://dx.doi.org/10.5018/economics-ejournal.ja.2014-4

Posel, D. R., & Casale, D. M. (2011). Relative standing and subjective well-being in south africa : The role of perceptions , expectations and income mobility. *Social Indicators Research*, *104*, 195–223.

Pouliakas, K., & Theodossiou, I. (2010, March). Differences in the job satisfaction of high-paid and low-paid workers across europe. *International Labour Review*, *149*(1), 1–29. Available from http://doi.wiley.com/10.1111/j.1564-913X.2010.00073.x

Pugno, M. (2007). 14 the subjective well-being paradox: a suggested solution based on relational goods. *Handbook on the Economics of Happiness*, 263.

Reichstein, T. (n.d.). Econometrics: Endogeneity. Available from http://www.soc.aau.dk/fileadmin/user_upload/kbm/VoF/Kurser/2011/kvantitative-metoder/Slides/TR-Endogeneity.pdf

Rhodes, S. R. (1983). Age-related differences in work attitudes and behavior: A review and conceptual analysis. *Psychological Bulletin*, *93*(2), 328–367. Available from http://content.apa.org/journals/bul/93/2/328

Roberts, J., Hodgson, R., & Dolan, P. (2011, September). "It's driving her mad": Gender differences in the effects of commuting on psychological health. *Journal of health economics*, *30*(5), 1064–76. Available from http://www.ncbi.nlm.nih.gov/pubmed/21855154

Roelen, C., Koopmans, P., & Groothoff, J. (2008, January). Which work factors determine job satisfaction? *Work (Reading, Mass.)*, *30*(4), 433–9. Available from http://www.ncbi.nlm.nih.gov/pubmed/18725706

Royo, M. G. (2007). Well-being and consumption: Towards a theoretical approach based on human needs satisfaction. *Handbook on the Economics of Happiness; Northampton, MA, US: Edward Elgar Publishing*, 151–169.

Sandvik, E., Diener, E., & Seidlitz, L. (1993). Subjective well-being: The convergence and stability of self-report and non-self-report measures. *Journal of Personality*, *61*, 317-342.

Sonnentag, S., Arbeus, H., Mahn, C., & Fritz, C. (2014, April). Exhaustion and lack of psychological detachment from work during off-job time: Moderator effects of time pressure and leisure experiences. *Journal of Occupational health Psychology*, *19*(2), 206–16. Available from http://www.ncbi.nlm.nih.gov/pubmed/24635737

Sousa-Poza, A., & Sousa-Poza, A. a. (2000, November). Well-being at work: A cross-national analysis of the levels and determinants of job satisfaction. *The Journal of Socio-Economics*, *29*(6), 517–538. Available from http://linkinghub.elsevier.com/retrieve/pii/S1053535700000858

Srivastava, S. (2013, July). Job satisfaction and organizational commitment relationship: Effect of personality variables. *Vision: The Journal of Business Perspective*, *17*(2), 159–167. Available from http://vis.sagepub.com/lookup/doi/10.1177/0972262912483529

Stanton-Rich, H. M., & Iso-Ahola, S. E. (1998, November). Burnout and leisure. *Journal of Applied Social Psychology*, *28*(21), 1931–1950. Available from http://doi.wiley.com/10.1111/j.1559-1816.1998.tb01354.x

Steinmetz, S., Vries, D. H. D., & Tijdens, K. G. (2014, January). Should I stay or should I go? The impact of working time and wages on retention in the health workforce. *Human Resources for Health*, *12*(1), 23. Available from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4021570&tool=pmcentrez&rendertype=abstract

Stock, J. H., & Watson, M. W. (2006). *Introduction to econometrics* (2nd ed.). Pearson.

Stock, J. H., Wright, J. H., & Yogo, M. (2002, October). A survey of weak instruments and weak identification in generalized method of moments. *Journal of Business & Economic Statistics*, *20*(4), 518–529. Available from http://www.tandfonline.com/doi/abs/10.1198/073500102288618658

Studenmund, A. H. (2005). *Using econometrics - a practical guide* (5th ed.). Addison Wesley.

Stutzer, A., & Frey, B. S. (2006, April). Does marriage make people happy, or do happy people get married? *The Journal of Socio-Economics*, *35*(2), 326–347. Available from http://linkinghub.elsevier.com/retrieve/pii/S1053535705001745

Stutzer, A., & Frey, B. S. (2008, June). Stress that doesn't pay: The commuting paradox. *Scandinavian Journal of Economics*, *110*(2), 339–366. Available from http://doi.wiley.com/10.1111/j.1467-9442.2008.00542.x

Stutzer, A., & Frey, B. S. (2010). Recent advances in the economics of individual subjective well-being. *Social Research: An International Quarterly*, *77*(2), 679–714.

Sutton, S. K., & Davidson, R. J. (1997). Prefrontal brain symmetry: A biological substrate of the behavioral approach and inhibition systems. *Psychological Science*, *8*(3), 204–210. Available from http://pss.sagepub.com/content/8/3/204.abstract

Templer, K. J. (2012, January). Five-factor model of personality and job satisfaction: The importance of agreeableness in a tight and collectivistic Asian society. *Applied Psychology*, *61*(1), 114–129. Available from http://doi.wiley.com/10.1111/j.1464-0597.2011.00459.x

Testing for nonspuriousness. (n.d.). Available from https://learn.bu.edu/bbcswebdav/pid-826908-dt-content-rid-2073693_1/courses/13sprgmetcj702_ol/week05/metcj702_W05S01T07_nonspur.html

Tiefenbach, T., & Kohlbacher, F. (2012, January). Subjektives glücksempfinden und seine einflussfactoren im japanischen kontext. *DEUTSCHES INSTITUT FÜR JAPANSTUDIEN, German Institute for Japanese Studies (DIJ)*, 1–43. Available from http://www.dijtokyo.org/publications/WP1201_Tiefenbach_Kohlbacher.pdf

Tokuda, Y., Hayano, K., Ozaki, M., Bito, S., Yanai, H., & Koizumi, S. (2009, April). The interrelationships between working conditions, job satisfaction, burnout and mental health among hospital physicians in Japan: a path analysis. *Industrial health*, *47*(2), 166–72. Available from http://www.ncbi.nlm.nih.gov/pubmed/19367046

Tourigny, L., Baba, V., & Wang, X. (2010, December). Burnout and depression among nurses in Japan and China: The moderating effects of job satisfaction and absence. *The International Journal of Human Resource Management*, *21*(15), 2741–2761. Available from http://www.tandfonline.com/doi/abs/10.1080/09585192.2010.528656

Trochim, W. M. (2006). Measurement error. Available from http://www.socialresearchmethods.net/kb/measerr.php

Varce, P. E., Shaffer, G. S., & Cynthia, M. (1983). Sex differences in job satisfaction revisited. *The Academy of Management Journal*, *26*(2), 348–353. Available from http://www.jstor.org/stable/255982

Veenhoven, R. (1991). Is happiness relative? *Social Indicators Research*, *24*(1), 1–34. Available from http://link.springer.com/article/10.1007/BF00292648

Veenhoven, R. (1994, June). Is happiness a trait? *Social Indicators Research*, *32*(2), 101–160. Available from http://link.springer.com/10.1007/BF01078732

Veenhoven, R. (1996, January). Developments in satisfaction-research. *Social Indicators Research*, *37*(1), 1–46. Available from http://link.springer.com/10.1007/BF00300268

Veenhoven, R. (1997). Advances in understanding happiness. *Revue Québécoise de Psychologie*, *18*, 29–74.

Vendrik, M., & Hirata, J. (2007). 10. Experienced versus decision utility of income: relative or absolute happiness. *Handbook on the economics of happiness*, 185.

Weaver, C. N. (1974). Correlates of job satisfaction: Some evidence from the national surveys. *The Academy of Management Journal*, *17*(2), 373–375.

Weaver, C. N. (1978). Sex differences in the determinants of job satisfaction. *The Academy of Management Journal*, *21*(2), 265–274. Available from http://www.jstor.org/stable/255759

Wooldridge, J. M. (2002). *Introductory econometrics*. South-Western.

Wright, J. D., & Hamilton, R. F. (1978). Work satisfaction and age: Some evidence for the job change hypothesis. *Social Forces*, *56*, 1140–1158.

Yamaguchi, I. (2013, January). A japan–us cross-cultural study of relationships among team autonomy, organizational social capital, job satisfaction, and organizational commitment. *International Journal of Intercultural Relations*, *37*(1), 58–71. Available from http://linkinghub.elsevier.com/retrieve/pii/S0147176712000569

Yarnal, C. M., Almeida, D. M., & Qian, X. L. (2014). Does leisure time moderate or mediate the effect of daily stress on positve affect? *Journal of Leisure Research*, *46*(1), 106–124. Available from http://www.midus.wisc.edu/findings/pdfs/1269.pdf

Zhai, Q., Willis, M., O'Shea, B., Zhai, Y., & Yang, Y. (2013, January). Big five personality traits, job satisfaction and subjective wellbeing in China. *International Journal of Psychology: Journal International de Psychologie*, *48*(6), 1099–108. Available from http://www.ncbi.nlm.nih.gov/pubmed/23075216

Statutory Declaration

I herewith declare that I have completed the present thesis independently making use only of the specified literature and aids. Sentences or parts of sentences quoted literally are marked as quotations; identifications of other references with regard to the statement and scope of the work is quoted. The thesis in this form or in any other form has not been submitted to an examination body and has not been published.