Three Essays in Applied Microeconomics

by

Xuelian Zhang

A Dissertation submitted to the

Graduate School-New Brunswick

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Economics

written under the direction of

Carolyn Moehling

and approved by

____

____

____

____

New Brunswick, New Jersey

May, 2013

ABSTRACT OF THE DISSERTATION

THREE ESSAYS IN APPLIED MICROECONOMICS

By Xuelian Zhang

Dissertation Director:

Carolyn Moehling

This dissertation contains three chapters in applied microeconomics. All three chapters try to answer one question: what factors determine labor market outcomes like employment probability, occupational choice and earnings?

Chapter 1 investigates the effects of multidimensional personality traits on employment status, occupational choice and earnings. Using the United Kingdom National Child Development Study, the analysis deals with the problems of reverse causality and measurement error by instrumental variable methods. The results indicate that personality traits play an important role in explaining the variation in labor market outcomes. The more agreeable and conscientious, and the less imaginative a person is, the more likely he is employed. The more outgoing and the less imaginative a person is, the more likely he works in a managerial occupation, but the less likely in a non-manual occupation. Agreeableness reduces one's probability of being in a professional occupation. Being outgoing and conscientious leads to higher earnings for paid employees.

Chapter 2 uses the United States Health and Retirement Study to study the effects of elder care provision on one's job choice with respect to flexibility. Fixed effects panel data models are used to control for time-invariant individual heterogeneity. Compared to non-caregivers, both male and female caregivers are significantly more likely to sort into flexible jobs or occupations, though they realize job flexibility through different channels: caregiving women are more likely to choose jobs with direct flexible work arrangements like flexible schedules, while caregiving men are more likely to realize flexibility indirectly by sorting into flexible occupation categories.

Chapter 3 uses the Brazil Living Standards Measurement Study Survey to examine the long-run consequences of child labor on an adult's income, health and educational attainment. The analysis leads to the following conclusions. Early working has a substantial negative impact on earnings for rural residents but no impact on urban residents. For health, child labor has an adverse consequence in the long run. As for the schooling effect, the earlier one enters the labor market, the fewer years of schooling he obtains. I also discover appreciable differences of child labor effects between urban and rural residents.

Acknowledgements

A great many people have contributed to the production of this dissertation. I would like to express my gratitude to all those people who have made this dissertation possible.

I owe sincere and earnest thankfulness to my advisor, Professor Carolyn Moehling. This dissertation would not have been possible without her guidance, stimulating suggestions and continuous support. Her encouragement and patience helped me overcome many difficulties. In addition to the guidance on research, Professor Moehling also helped me in many ways, like helping me improve interpersonal skills and presentation skills. For everything you have done for me, Professor Moehling, I thank you.

I am deeply grateful to Professor Roger Klein, Professor Mark Killingsworth and Professor John Bowblis for their insightful comments and suggestions on my dissertation. I would like to thank them for their time and efforts to serve on my committee.

My thanks also go to the staffs in the Department of Economics at Rutgers University, for their various forms of support during my graduate study.

Many friends have helped me stay sane through these difficult years. I would like to thank their friendship and love.

Most importantly, I would like to give my special thanks to my husband Dungang Liu and my family in China whose love continuously supports me and encourages me in all these years.

Table of Contents

List of Tables

List of Illustrations

Part I

Introduction

A key question in labor economics is: what factors determine labor market outcomes like employment probability, occupational choice and earnings? The existing literature mainly focuses on factors like demographics, education and work experience. However, these factors can only account for a small portion of the total variation in labor market outcomes (Osborne Groves, 2005). There should exist many other factors which have important explanatory power for labor market outcomes but are not well-examined. In this dissertation, I investigate three factors not included in traditional examinations: personality traits, early work in childhood, and the responsibility in caring for elderly parents. Each of these factors raises many important questions that are of great interest in modern society. For example, in developed countries, how can personality traits or elder care responsibilities impact an individual's occupational choice? In developing countries, does early work in childhood always have adverse effects on an individual's development? This dissertation establishes the effects of these factors on an individual's labor market outcomes. It also discusses the public and social implications of the findings.

1 Do Personality Traits Matter in the Labor Market? Evidence from the United Kingdom

The relationships between one's personality traits and labor market outcomes has become a hot topic. Due to the availability of personality measures in large survey data sets in recent years, studies have demonstrated the strong correlations between personality traits and labor market outcomes, with analyses from the United States (Osborne Groves, 2005; Heckman, Stixrud and Urzua, 2006; Mueller and Plug, 2006), the United Kingdom (Fronstin, Greenberg and Robins, 2005; Jackson, 2006; Nandi and Nicoletti, 2009), Germany (Heineck and Anger, 2010) and some other countries (Nyhus and Pons, 2005; Semykina and Linz, 2007). A more challenging question is whether personality traits causally affect labor market outcomes. One difficulty in estimating such causal effects is the possible existence of reverse causality, i.e., labor market outcomes may reversely affect an individual's personality traits. Studies have shown that one's personality traits can be significantly affected by economic outcomes (see for example, Gottschalk, 2005 and Powdthavee, Boyce and Wood, 2011). Therefore, while many of the existing studies simply treat personality traits as exogenous to labor market outcomes, they can only provide suggestive correlations between personality traits and outcomes.

This chapter contributes to the literature by addressing this issue and estimating the causal effects of the Big Five personality traits on employment probability, occupational choice and earnings. The Big Five personality traits are extraversion, agreeableness, conscientiousness, emotional stability and imagination. Psychologists characterize the Big Five as the "latitude and longitude" of personality traits and argue that most personality constructs can be mapped onto these five traits (Costa and McCrae, 1992a; Goldberg, 1993; Ozer and Raise, 1994). Using the United Kingdom National Child Development Study, I address the problem of reverse causality using an instrumental variable method, where adulthood personality traits are instrumented by proxies of childhood personality traits. The instrumental variable method also mitigates the possible measurement error problem arising from the use of self-reported subjective personality scales.

I find that personality traits play an important role in explaining the variation in labor market outcomes. For example, the results suggest an approximately 6 and 8 percentage points higher probability of being employed for an individual by a one standard deviation increase in agreeableness and conscientiousness respectively, but a 26 percentage points lower probability of being employed by a one standard deviation increase in imagination. The negative relationship between imagination and employment probability seems surprising, but it may be a reflection of a positive correlation between imagination and non-conformity, a trait not valued by most employers. Moreover, I find that an individual is more likely to work in a managerial occupation but less likely to work in a non-manual occupation if he is more outgoing and less imaginative. Agreeableness reduces one's probability of being in a professional occupation. Furthermore, being outgoing and conscientious lead to higher earnings for paid employees.

These findings have important policy implications. For instance, education system and worker training programs may consider incorporating the training of behavioral and social skills, in addition to cognitive skills.[1] This could help individuals become more competitive and be more likely to find employment in the labor market. Also, the results in this chapter can provide useful guidelines for job counselors on how to help workers with different personality traits find the occupations that are most suitable for them.

2 Do Elder Care Providers Sort into Flexible Jobs? Evidence from the Health and Retirement Study

Due to the aging of baby boomers, the elderly population in the US has increased from 35 million to 40 million in the past decade, and it is expected to further increase to 72.1 million by 2030 (U.S. Department of Health and Human Services, 2011). This rapid increase in the elderly population has raised many important social concerns, one of which is the heavy burden of elder care. Currently, working-age adult children are common sources of care provision. The competing time demands of work and care impose great challenges. Studies have found that care can significantly reduce paid employment by forcing care providers to work fewer hours or even totally withdraw from the labor market (see for example, Ettner, 1995, 1996), and at the same time employment can negatively affect the likelihood of providing care (Boaz and Muller, 1992; Michaud, Heitmueller and Nazarov, 2010). Therefore, balancing work and family is a major issue for care providers.

Flexible work arrangements (e.g. flexible schedule) are frequently proposed as a way to accommodate the needs of working care providers (see for example, Heitmueller, 2007; Bolin, Lindgren and Lundborg, 2008). This raises an interesting question: compared to non-caregivers, are caregivers more likely to sort into jobs with flexible work arrangements? Ex ante, it is difficult to answer this question due to the possible two-fold consequences of flexible work arrangements. On the one hand, flexible work arrangements can help reconcile work and care, and thus generate less pressured working conditions for care providers. On the other hand, they are often associated with some negative consequences, like lower wages, reduced promotion opportunities, and so on (Rhoads, 1993, P18; Heywood, Sieberty and Weiz, 2007). As a result, it remains an open question as to whether caregivers consciously

[1] A similar implication is also mentioned by Nyhus and Pons (2005) and Osborne Groves (2005).

sort into flexible jobs or not.

Using the national representative longitudinal data from the United States Health and Retirement Study (HRS), I analyze whether the provision of care affects an individual's choice of a flexible job. However, there exists an empirical challenge caused by the existence of individual heterogeneity. An individual's unobserved time-invariant characteristics (e.g. personality) may affect both job choice and care provision, and this will bias the estimation of the effect of care provision on an individual's flexible job choice. In this chapter, time-invariant individual heterogeneity is controlled by using fixed effects panel data models.

I find that, for both men and women, care responsibility causes individuals to be significantly more likely to sort into flexible jobs or occupations, compared to non-caregivers. This result is robust to different measures of job flexibility and different care definitions. However, men and women realize job flexibility through different channels: caregiving women are more likely to directly choose jobs with flexible work arrangements like flexible schedules, while caregiving men are more likely to realize flexibility indirectly by sorting into flexible occupation categories. The results suggest that for care providers, the need to balance work and family outweighs the possible negative consequences like lower wages associated with flexible work arrangements. Therefore, workplaces may want to provide more flexible work arrangements to better accommodate caregivers' needs for balancing paid employment and unpaid care work.

3 Long-term Health and Socioeconomic Consequences of Child Labor: Evidence from Brazil

Child labor has long been considered as a social problem in developing countries. However, there is an ongoing debate on whether child labor always plays a negative role in children's development (see for example the discussions in O'Donnell, Rosati and Doorslaer (2005) and Emerson and Souza (2007)). Therefore, it is important to understand how child labor can affect adult outcomes, such as earnings, health, and so on.

This chapter investigates multiple adult outcomes of child labor simultaneously. These outcomes include adult earnings, health and schooling. The analysis is based on the Brazil Living Standards Measurement Study Survey. A challenging problem here is that there may exist some unobserved factors (e.g. ability) correlated to both early working and adult outcomes of an individual. This can lead to estimation bias of the effect of interest, so I use an instrumental variable method to correct such bias.

The analysis leads to the following conclusions. With respect to earnings, early working has a substantial negative impact on rural residents but no impact on urban residents, after controlling for schooling and health status. For health, child labor has an adverse consequence in the long run. As for the schooling effect, the earlier one enters the labor market, the fewer years of schooling s/he obtains. I also discover appreciable differences between urban and rural residents. For example, rural residents suffer a greater adverse health effect of early working than those in urban areas. These differences should be taken into account when child labor policies are proposed. These findings have many important implications. Overall these results make a strong call to reduce child labor in Brazil and other developing countries. Reducing workloads in childhood leads to higher income, better health and increased schooling.

Part II

Chapter 1

Do Personality Traits Matter in the Labor Market? Evidence from the United Kingdom

1.1 Introduction

It is very common to treat cognitive skills (often represented by test scores) as important determinants of labor market outcomes like earnings, employment status and occupational choice. However, as Nyhus and Pons (2005) and Osborne Groves (2005) points out, still much of the variation in labor market outcomes remains unexplained. For instance, only 10% to 40% of the variation in earnings can be explained by traditional cognitive skills and other human capital and background variables (Osborne Groves, 2005).

In order to better understand the determinants of labor market outcomes, economists are now exploring the role of personality traits. An individual's personality traits, such as gregariousness, self-organization and emotional stability, should affect his behavior and matter for economic success. However, personality measures have not been included in large survey datasets until recent years. Now there is a growing body of research that examines the correlations between personality traits and labor market outcomes, with data from the United States (Osborne Groves, 2005; Heckman, Stixrud and Urzua, 2006; Mueller and Plug, 2006; Drago, 2011; Fletcher, 2012), the United Kingdom (Fronstin, Greenberg and Robins, 2005; Jackson, 2006; Nandi and Nicoletti, 2009), Germany (Heineck and Anger, 2010) and some other countries (Nyhus and Pons, 2005; Semykina and Linz, 2007).

However, many of the existing studies simply treat personality traits as exogenous to labor market outcomes. They do not address the possible issue of reverse causality, i.e., labor market outcomes may influence personality traits. Some researchers argue that personality traits are stable after early adulthood (Costa and McCrae, 1994; McCrae and Costa, 1994), and therefore are not affected by social environment, thus there is no harm to treat personality traits as exogenous. However, recent studies suggest that personality characteristics are not stable and could have non-negligible changes after early adulthood (Srivastava et al., 2003; Roberts, Walton and Viechtbauer, 2006). Furthermore, several studies establish that personality can be significantly affected by economic outcomes (Gottschalk, 2005; Sutin et al., 2009; Powdthavee, Boyce and Wood, 2011). Therefore, it may not be appropriate to treat personality traits as stable and exogenous when examining their effects on labor market outcomes. Those studies which do not address potential reverse causality can only provide suggestive correlations between personality traits and outcomes (see Almlund et al. (2011) for more detailed discussions about the possible issue of reverse causality and the challenges one may face when treating personality traits as exogenous).

Correlation does not imply causation. Identifying the causal mechanism underlying the relationships between personality traits and labor market outcomes is of great importance. If personality traits strongly influence an individual's labor market success, this would indicate the importance of personality traits in explaining the variation in labor market outcomes. Also, we need causal relationships rather than correlations between personality traits and labor market outcomes for the purpose of policy guidance. In addition, the causal linkages between personality traits and occupational choice could provide useful guidelines for job counselors for helping workers with different personality traits find the occupations that are most suitable for them.

This chapter investigates the causal effects of the multidimensional Big Five personality traits on an individual's employment status, occupational choice and earnings. The Big Five personality traits provide a comprehensive picture of an individual's personality profile. Specifically, the Big Five personality traits include extraversion, agreeableness, conscientiousness, emotional stability and imagination.[1] The present study adds to the literature by examining the causal relationships between the Big Five personality traits and labor market outcomes with instrumental variable techniques. Using the National Child Development Study (NCDS) from the United Kingdom (University of London, 2008; University of London, 2010), this chapter employs teacher's and mother's assessments of the respondent's childhood behavior and social adjustment to instrument for adult personality traits. Here, social adjustment means the adaptation of a person to the social environ-

[1] Further details for the Big Five personality traits are provided in Section 1.2.

ment (Campbell, 2004). Childhood behavior and social adjustment are good proxies for childhood personality, and studies have shown a continuity of personality from childhood to adulthood (Caspi, 2000). Therefore, the instruments and adult personality traits are closely correlated. After controlling for adult personality, the pre-market nature of childhood behavior and social adjustment prevents the possible correlations between the instruments and unexplained component of labor market outcomes, which satisfies the exclusion restriction for instruments. Another great advantage of instrumental variable techniques is mitigating the measurement error problem. As Bertrand and Mullainathan (2001) point out, an individual's response to subjective questions may be greatly affected by many factors: the ordering of questions, the social nature of the survey, question wording, etc. Also, self-reported responses may be correlated with a respondent's motivation, education level and some other factors. Therefore, there may exist great measurement errors in the self-reported subjective personality scales which will cause bias to estimates, and instrumental variable techniques could help to reduce such bias.

1.2 Background and Previous Findings

Although studies documenting the role of personality traits in the labor market are limited, research in this area has expanded rapidly in recent years due to the availability of personality measures in large survey data sets. A brief review of the recent literature helps us identify what scholars have already discovered and what still remains to be explored.

As mentioned above, when examining the causal effects of personality traits on labor market outcomes, it is not appropriate to treat personality traits as exogenous. Studies have shown the reverse effects of economic outcomes on personality characteristics. For instance, Sutin et al. (2009) find that an individual's higher income prospectively predicts an increase in emotional stability one decade later. Using surveys of lottery winners from the UK, Powdthavee, Boyce and Wood (2011) suggest that an increase in unearned income causes people to become more gregarious, sympathetic and emotionally stable. Gottschalk (2005) uses data from a randomized control trial, and documents the significant influence of working at a job on people's belief of their control over things that happen to them.

Therefore, we cannot ignore the issue of reverse causality if we attempt to investigate the causal effects of personality traits on labor market outcomes. In order to address this issue, some scholars employ early-stage measures of personality to predict later labor market outcomes. In this way, they argue that there no longer exists endogeneity problem (e.g., Osborne Groves, 2005; Jackson, 2006; Carneiro, Crawford and Goodman, 2007). For example, based on the examination of the NCDS from the UK, Jackson (2006) finds that the syndromes of aggression and withdrawal exhibited at childhood adversely affect the probability of being employed and occupational attainment in adulthood. Here, the aggression syndrome refers to "the anxious, aggressive, restless, outwardly expressed behavior", and the withdrawal syndrome refers to "the anxious, withdrawn inhibited behavior" (Ghodsian, 1977). For each additional point in the withdrawal score (the scale for the withdrawal score is 0 to 4.9), an individual's chance of entering the managerial class in adulthood is reduced by 22% compared to the chance of entering the working class. In addition, aggression imposes a significant negative effect on entry to the higher technical class in adulthood. Osborne Groves (2005) shows that, among participants of the National Longitudinal Survey of Young Women (NLSYW), childhood locus of control adversely affects female wages. While persons with an internal locus of control believe that their own decisions and behavior determine life events, those with an external locus of control believe that powerful others, fate or luck rather than themselves primarily determine their life events (Rotter, 1966), Osborne Groves (2005) finds that a one standard deviation increase in the Rotter Locus of Control score decreases adult wages by 5.5%. In the same study using data from the NCDS, she shows that childhood aggression and withdrawal syndromes significantly reduce wages for female women in the UK. Duncan and Dunifon (1998) show that motivation scales measured around 1970 impose non-negligible impacts on later hourly earnings measured around 1990, based on the Panel Study of Income Dynamics.

However, as Jackson (2006) and Almlund et al. (2011) point out, when employing early-stage personality measures to predict later labor market outcomes, we may need to worry about the representativeness of the early-stage personality measures on the current personality traits. It is the current personality traits that cause outcomes, not the early personality traits. While personality traits can be affected by social environment and changes significantly over time (Helson, Jones and Kwan, 2002; Roberts, Walton and Viechtbauer, 2006), early-stage personality measures may be poor predictors for later outcomes. Therefore, employing early-stage personality measures would result in an errors in variables problem if personality traits change over the relevant time frame (Almlund et al., 2011).

Some other studies address the issue of reverse causality with different approaches. For instance, when analyzing the effect of self-esteem on earnings among participants of NLSY79, Drago (2011) deals with reverse causality by instrumenting a measure of self-esteem in 1987 with a measure of self-esteem in 1980. In order to examine the effects of locus of control on earnings for the NLSYW participants, Osborne Groves (2005) creates an instrument for adult personality by removing the effect of past wages on the adult locus of control scale. However, most of these studies use personality measures that only describe a single trait. There are many aspects of an individual's personality traits, like gregariousness, self-discipline and emotional stability. Hence, it may be too restrictive to employ a single trait to capture the impacts of personality traits on labor market outcomes. Research employing single-trait measures may not be able to provide a full picture of the labor market consequences of personality traits.

Contrary to those single-trait personality measures, the Big Five provides a comprehensive picture of an individual's multidimensional personality traits. Psychologists characterize the Big Five as the "latitude and longitude" of personality traits and argue that most personality constructs can be mapped onto the Big Five (Costa and McCrae, 1992a; Goldberg, 1993; Ozer and Raise, 1994). Specifically, the factors contained in the Big Five are extraversion, agreeableness, conscientiousness, emotional stability and imagination.[2] Extraversion refers to "the act, state, or habit of being predominantly concerned with and obtaining gratification from what is outside the self".[3] Extraverts are generally enthusiastic and talkative, and good at interacting with other people. Agreeable persons tend to be sympathetic, considerate and cooperative toward others.[4] They are usually concerned about pleasing others. Conscientiousness is related to an individual's diligence, self-discipline, self-organization and acceptance of responsibility.[5] Emotional stability incorporates the

[2] Sometimes "imagination" is replaced by other factors, like openness to experience (as used in the German Socio-Economic Panel Study), autonomy (as used in the DNB Household Survey), etc.

[3] Merriam Webster Dictionary.

[4] en.wikipedia.org/wiki/Agreeableness (accessed on 3/26/2013).

[5] en.wikipedia.org/wiki/Conscientiousness (accessed on 3/26/2013).

attributes of being calm and even-tempered.[6] People with emotional stability are usually good at dealing with stress and negative emotions, and do not have rapid mood changes. Imagination is also referred to as intellect[7], indicating a person is imaginative, creative and curious. The "imagination" scale in the International Personality Item Pool (IPIP), which is employed in the current study, is closely related to the "openness to experience" scale in the NEO Personality Inventory which is also widely used in psychology literature[8]. Although the creativity captured by the "imagination" scale may be rewarded in some fields (e.g., academia), as many scholars (see for example, Johnson, 1983; Judge et al., 1999; Heineck and Anger, 2010) point out, persons with high levels of imagination or openness are likely to be autonomous and nonconforming which may be of little help or even an obstacle to job market success.

In order to identify the relationships between multidimensional personality traits and economic success, many scholars have employed the Big Five to measure personality profile. Nyhus and Pons (2005) observe that emotional stability leads to higher wages for both men and women, but agreeableness is associated with lower wages for female workers in Netherlands. Based on the examination of the Wisconsin Longitudinal Study, Mueller and Plug (2006) document that antagonism (the opposite of agreeableness), emotional stability and openness to experience are rewarded in the labor market for men, while women who are conscientious and open to experience receive higher earnings. In their analysis of the German Socio-Economic Panel Study, Heineck and Anger (2010) employ error-in-variables (EIV) regressions to correct for the measurement error problem in the pooled cross-sectional estimation, and find that while there are wage premiums for males with higher levels of extraversion and conscientiousness, male workers who are open to experience are punished in the labor market by lower wages. In the same study, after controlling for individual heterogeneity with panel approaches, Heineck and Anger find that antagonistic women earn more, but none of the Big Five personality traits has a significant effect on men's

[6] en.wikipedia.org/wiki/Neuroticism (accessed on 3/26/2013).

[7] The present study uses "imagination" rather than "intellect" to avoid confusion between intellect personality and intelligence. According to John and Srivastava (1999), the intellect personality only has very small correlations with IQ measures, it is not a measure of intelligence.

[8] The details of the "imagination" scale and "openness to experience" scale can be obtained at http://ipip.ori.org/newMultipleconstructs.htm (Accessed on 9/3/2012).

wages. However, none of these studies solve the problem of reverse causality when examining the effects of the Big Five personality traits on labor market outcomes.[9] As there exists empirical evidence of the existence of reverse causality, ignoring this problem will cause bias to the estimates of the labor market effects of personality traits.

1.3 Methodology

The present study examines how the Big Five personality traits affect an individual's labor market outcomes, including employment status, occupational choice and earnings.

Different personality traits are valued differently in the labor market and thus persons with different personality traits may have different probabilities of being employed. For example, persons who are very diligent and organized may be more likely to be employed than persons who are relatively less conscientious. To examine whether and how an individual's personality traits affect his likelihood of being employed, the following probit model is employed:

$$employed = \begin{cases} 1 & \text{if employed*>0} \\ 0 & \text{otherwise} \end{cases}$$

$$\begin{aligned} employed_i^* &= \beta_0 + \beta_1 extraversion_i + \beta_2 agreeableness_i + \beta_3 conscientiousness_i \\ &\quad + \beta_4 emotionalstability_i + \beta_5 imagination_i + x_i^{'}\gamma + \varepsilon_i \qquad (1.1) \end{aligned}$$

where *employed* denotes the likelihood of being employed and *employed** is the latent variable measuring the propensity of being employed. Here, x is a set of observed individual characteristics other than the Big Five personality traits.

Personality traits may also predict sorting into occupations. From an economic perspective, individual differences in personality traits may be linked to different abilities and

[9] Nyhus and Pons (2005) regress the personality traits on age and use the predicted residuals as measures of personality to remove the effects of age on personality. Similar to Nyhus and Pons (2005), Heineck and Anger (2010) also remove the age effects by regressing personality traits on age and age squared and using the residuals as personality measures in their main context analysis. They argue that this way could somewhat pick up the possible reverse effect of job environment on personality. In addition, as a robustness check, Heineck and Anger regress wages on lagged measures of personality to mitigate reverse causality. However, only employing age-effect-free personality traits is not enough to solve the reverse causality problem, and we may need to worry about the representativeness of lagged personality measures when using them as proxies for current personality traits.

preferences which are important determinants of occupational choice. In order to check whether this is the case, I employ the following multinomial logit (MNL) model to estimate the probability of individual i being observed in occupation j:

$$\Pr(Occupation_i = j \mid w_i) = \frac{\exp(w_i^{'}\alpha_j)}{\sum_{j=1}^{4}\exp(w_i^{'}\alpha_j)}, \qquad j = 1, 2, 3, 4 \tag{1.2}$$

Occupation represents an individual's occupational choice, w includes the Big Five personality traits and the set of other individual characteristics x which is the same to that in the employment status model. This study divides the occupations into four categories: managerial, non-manual and technical (abbreviated to non-manual), professional and manual. Here, the coding of managerial and professional occupation categories is guided by Jackson (2006), and the occupations other than managerial and professional are aggregated and recoded into manual and non-manual categories.[10] Managerial occupation category mainly includes manager and employer, professional category includes lawyer, engineer, actuary, and the like, non-manual category includes artist, author, photographer, draughts person, etc., and manual category includes van driver, printer, baker and so on. The descriptions and some examples corresponding to each occupation category are shown in Table A.1.1 in the Appendix.

In order to investigate the effects of personality characteristics on earnings, the following linear model is employed:

$$\begin{aligned} \ln(earnings_i) \quad &= \quad \delta_0 + \delta_1 extraversion_i + \delta_2 agreeableness_i + \delta_3 conscientiousness_i \\ &\quad + \delta_4 emotionalstability_i + \delta_5 imagination_i + x_i^{'}\theta + \varepsilon_i \end{aligned} \tag{1.3}$$

where ln*($earnings_i$)* is the log of weekly earnings on the current main job, x consists of the same set of individual characteristics as in the employment status model.

However, the five personality measures may be correlated with the unexplained component of labor market outcomes due to reverse causality and measurement error, and

[10] Due to the limited sample size in this model, although the occupation categories for "manual" and "non-manual" are broad, they cannot be divided into more detailed categories.

this will bias the estimates in the above models. Success or failure in the labor market may affect a person's personality traits. For example, unfavorable work experiences may make an individual feel upset, irritable and emotionally unstable. Being unemployed for a long time may make an individual feel too embarrassed to interact with other persons and thus reduces his extraversion. The competition and stress in working may cause workers to become more antagonistic and less agreeable. However, labor market success may free people from financial obligations and thus make them more likely to be gregarious, emotionally stable and open to new experience. In addition, the self-reported subjective personality measures may be subject to large measurement error. This problem arises from the fact that the questions designed to measure an individual's personality traits are only imperfect proxies of the true traits. Also, as Bertrand and Mullainathan (2001) suggest, an individual's response to subjective questions may be greatly affected by many factors: the ordering of questions, the social nature of the survey, question wording, etc. All these may lead to measurement error in subjective measures. In addition, an individual's response to subjective questions may be affected by his own values. For example, if a person feels that conscientiousness is a valuable trait, he may intentionally increase self-assessment on his measure of conscientiousness, and this would lead to measurement error.

In order to address the issues of reverse causality and measurement error, I use instrumental variable methods. Teacher's and mother's assessments of the respondent's childhood behavior and social adjustment serve as instruments for adult personality traits. Consequently, an IVprobit specification is employed for the employment status model (1.1), and a two-stage least squares (2SLS) specification is employed for the earnings model (1.3). For the occupational choice model (1.2), a two-stage multinomial logit (2SMNL) model is employed: regress the five endogenous personality traits on the instruments and exogenous variables (x) in the first stage with OLS, get predicting values for the five personality traits, and then regress occupational choices on the five predicted personality traits and exogenous variables (x) with MNL in the second stage.

1.4 Data

1.4.1 Primary Data

The data come from the National Child Development Study (NCDS), a nationally representative longitudinal study following the lives of all individuals who were born in England, Wales or Scotland between the 3rd and 9th of March, 1958. There have been 9 waves to date, the first of which was conducted at birth, with follow-ups at ages 7 (1965), 11 (1969), 16 (1974), 23 (1981), 33 (1991), 42 (2000), 46 (2004) and 50 (2008). The NCDS is an extremely rich dataset, providing detailed information about the respondent's life, including labor market outcomes, psychological characteristics, educational histories, family background and some other aspects. The unique feature of the NCDS is that it provides psychological information for the respondent both in his/her childhood and adulthood, and thereby fits my research purpose.

Regarding adult personality traits, in the latest wave of NCDS (2008), respondents were asked to self-rate their personality characteristics based on 50 questions from the International Personality Item Pool (IPIP) (Goldberg, 1999). These 50 questions were assigned to the Big Five personality traits, i.e., extraversion, agreeableness, conscientiousness, emotional stability and imagination. To provide a general idea about these five personality traits, I list the questions belonging to each trait in Table A.1.2 at Appendix. Scores on each personality trait were computed from the responses, with higher scores implying higher levels of that trait. The distributions of the five personality factors are presented in Figure 1.1.[11] It is easy to find that the majority of the respondents report relatively high scores. I also provide the correlations between personality traits in Table A.1.3 at Appendix. It is clear that the Big Five personality traits are not quite correlated with each other.

The present study analyzes how personality traits would help to explain the differentials in employment status, occupational choice and earnings for individuals at age 50. Due to the distinct labor market behavior of male and female workers, the sample is restricted to males with valid information on personality traits when examining the influences of personality traits on employment status.[12] The analysis on how personality characteristics affect an individual's occupational choice is restricted to male workers with valid occupation and

[11] The distributions of the five personality traits are based on the selected male sample, and the sample selection process is discussed later.

[12] For future research, examining the labor market consequences of personality traits for female workers could be worthy of study, though such analysis may be complicated by fertility, child care and other decisions of women.

personality information. Lastly, when exploring the effects of personality traits on earnings, the sample is further restricted to full-time male workers with valid information on earnings and personality traits. For all the three models, the samples include both self-employed workers and paid employees.[13] After the selection process, the sample size is 2105 for the employment status model, 1928 for the occupational choice model and 1618 for the earnings model.[14]

While there were 4822 males interviewed in the 2008 wave of NCDS, we may need to worry about the representativeness of the selected sample. I compare the means of the variables used in the analysis for the selected sample and for those out-of-sample observations which are in the original data set. Some statistically significant differences exist: the individuals in the selected sample have higher employment probability in general, higher probability in non-manual occupations and lower probability in manual occupations, as well as higher earnings. In-sample individuals also have higher education levels and reading test scores at age 11. However, there do not exist significant differences in adult personality traits. Therefore, we may expect a bias on the effects of personality traits on labor market outcomes when we focus on the selected sample. Nevertheless, although these differences are statistically significant, they are small in magnitude. For instance, 92% of the selected sample and 88% of the out-of-sample observations are employed, 27% of the selected sample and 24% of the out-of-sample population works in non-manual occupations, and while the average log of weekly earnings for in-sample individuals is 6.47, the corresponding value for the out-of-sample individuals is 6.37. Moreover, the distributions of the variables in the analysis are quite similar for the in-sample and out-of-sample observations. As a result, the possible bias in estimating the impacts of adult personality traits on labor market outcomes should be small. This is further reinforced by robustness checks.[15] In addition, I will discuss

[13] As an extension to the analysis in the main context, I re-estimate all models with the samples excluding self-employed workers in the Extension Section, and compare the results to those gotten in the main context analysis.

[14] The sample size is greatly reduced due to missing or incomplete information on the key variables: over 70% of the dropped observations are due to missing or incomplete information on the Big Five personality traits or childhood behavior assessments, and the rest dropped observations are due to missing information on dependent variables or other explanatory variables.

[15] I call "Full Sample 1" for all the observations with non-missing adult personality measures (note that these observations may have missing values on the childhood behavior assessments). I run a probit model for employment status, a multinomial logit model for occupational choice and an OLS model for earnings on the adult personality traits and other explanatory variables, with the Full Sample 1 and the selected sample being

later about the possible selection bias in the earnings model when the sample is restricted to the full-time workers with valid earnings, since the direction of the bias depends on the effects of personality traits on employment probability.

The summary information for the variables in the three models is presented in Table 1.1. The first two columns (Sample 1) exhibit the summary information for the sample of the employment status model, the second two columns (Sample 2) are for the sample of the occupational choice model, and the last two columns (Sample 3) show the information for the earnings model. Around 92% of males are employed, and the average of the log weekly earnings of male full-time workers is 6.47.[16] For the adult personality traits which range between 10 and 50 points, males rate themselves at relatively high scores on average, with 32 as the lowest for extraversion and 38 as the highest for agreeableness. More than 70% of the population is married, the same fraction lives in urban areas, and about 5% is from Wales, 12% from Scotland and 83% from England. For the education levels, around 49% of the population completed compulsory education, 9% completed extended secondary education, over 25% completed undergraduate or postgraduate education, and less than 17% of the population did not get any academic qualification.[17] While the maximum score for the reading test taken by the respondent at age 11 is 35 points, the average score for this test is only half of the maximum points. Fathers of the respondents took around 10 years of schooling on average.

1.4.2 The Instruments

As mentioned in the Methodology Section, instrumental variable methods are used to

used respectively, to examine the associations between adult personality traits and labor market outcomes. No statistically significant differences exist for the estimates of interest based on the two samples.

Full Sample 2 consists of all the observations with non-missing information on childhood behavior assessments (note that these observations may have missing values on adult personality measures). Similarly, I employed a probit model for employment status, a multinomial logit model for occupational choice and an OLS model for earnings on the childhood behavior assessments and other explanatory variables, to examine the associations between childhood behavior and adult labor market outcomes, based on Full Sample 2 and the selected sample respectively. Again, no statistically significant differences exist for the estimates of interest based on the two samples.

[16] The average of the weekly earnings of male full-time workers is 975.42 pounds.

[17] Children in England are required to take the compulsory education between age 5 and 16. At school year 11, they need to take certain exam (e.g. O-level exam) marking the end of compulsory education. After finishing compulsory education, some children choose to continue their secondary studies for a further one or two years, leading most typically to A-level qualifications or some other extended secondary education qualifications. Wales and Scotland have comparable academic qualification systems to England.

address the issues of reverse causality and measurement error. To be valid, instruments must be closely correlated with the endogenous variables and uncorrelated with the unexplained component of labor market outcomes. A candidate set of instruments arises from the survey questions answered by the respondent's mother and teacher about his behavior and social adjustment in childhood. In 1965, mothers were asked whether, and the extent to which, the children exhibited some particular kinds of behavior at age 7, like preferring to do things on his own rather than with others, worrying about many things, being easy to get upset by new situation, and so on. In 1969, teachers were asked to complete the Bristol Social Adjustment Guide (BSAG) for the children when they were 11 years old. The BSAG is a standardized psychological test of social adjustment. It consists of 146 items, each of which can be designated as one of the 12 domains (hostility towards adults, immature behavior, nervous symptoms, etc.). Teachers were asked to underline the items that they thought apply to the child, and each underlined item contributed a score of 1, from which scores for 12 domains of social adjustment were computed. The higher score a child gets for a domain, the more maladjusted he is in that aspect.

This research employs selective mother's and teacher's assessments of childhood behavior and social adjustment as the instruments for adult personality traits. To be more specific, the instruments include teacher's ratings about the child's hostility towards adults, immature behavior and nervous symptoms at age 11,[18] as well as mother's judgments about whether the child had the following behavior at age 7: having difficulty in settling to anything for more than a few moments, preferring to do things on his/her own rather than with others, being bullied by other children, being squirmy or fidgety, worrying about many things, being irritable or quick to fly off the handle, being upset by new situation or things happened for the first time, biting nails, and being disobedient at home. The distributions of teacher's ratings about child behavior are presented at Figure 1.2.[19] We find that most children got low scores for the three teachers' ratings.

[18] "Immature behavior" and "nervous symptoms" are originally labeled as "miscellaneous symptoms" and "miscellaneous nervous symptoms" in the BSAG. "Miscellaneous symptoms" is relabeled since they mainly refer to immaturity in the BSAG. Similarly, "miscellaneous nervous symptoms" is relabeled since they mainly refer to nervous behavior.

[19] The distributions of mother's judgments about childhood behavior are not presented, since all of them are dummies.

I employ both mother's and teacher's assessments of childhood behavior and social adjustment to work as the instruments for adult personality, because children may behave differently at school and at home. Also, the use of both mother's and teacher's assessments could provide a more complete picture of child behavior: mothers may observe their children more closely and carefully than teachers, while teachers may provide more objective ratings for children's behavior (Fronstin, Greenberg and Robins, 2005).

Teacher's and mother's assessments of childhood behavior and social adjustment are closely related to child personality and consequently adult personality. The exclusion restriction that the instruments should be uncorrelated with the unexplained component of employment status is automatically satisfied, owing to the pre-market nature of the assessments of childhood behavior and social adjustment, once adult personality traits are controlled for.[20]

When we look at the summary statistics for instruments (Table 1.1), we find that teachers generally assigned low scores for children's behavior at their age 11. For example, while the score for hostility towards adults range between 0 and 15 points, the average score that children got is only 0.7739 points, indicating that teachers did not think children were very hostile towards adults. Similarly, the average scores for immaturity and nervous symptoms are low, indicating that children on average were not immature or nervous. When we look at the mothers' judgments on whether children had certain behavior at age 7, we see more variations. For example, around 31% of children had difficulty in concentrating, 67% of them preferred to do things on their own, and more than 40% of children were fidgety, worried about many things or were irritable.

1.5 Empirical Results

1.5.1 Personality and Employment Status

In order to estimate the effects of personality traits on employment status, I begin by treating the five personality traits as exogenous. The control variables include marital sta-

[20] There may exist a concern about the exogeneity of the instruments, as parents or teachers may affect a child's career path based on their assessments of childhood behavior, and consequently have an influence on the child's adult labor market outcomes. However, I believe that the time distance between age 7(11) and age 50 is so long that the possible influences of parents or teachers based on childhood behavior at age 7 and 11 should be too small to affect an individual's adult labor market outcomes at age 50.

tus, current residence (urban/rural), region dummies for Wales and Scotland (the reference group is England), dummies for the highest educational attainment (the base group is "the individual did not finish secondary education"), the reading test score at age 11 which is a proxy for cognitive skills, and father's years of schooling which is a proxy for family background.[21] The first two columns of Table 1.2 report the results for the probit model (1.1). Here, the marginal effects rather than probit estimates are reported. Four out of the five personality traits are correlated with an individual's likelihood of being employed, but at very small magnitudes: a one standard deviation increase in agreeableness and imagination is associated with a decrease in employment probability of approximately 1 and 3 percentage point respectively, and when there is a one standard deviation increase in conscientiousness or emotional stability, the individual's probability of being employed would increase by around 2 percentage points.[22] Compared to the persons who do not have any academic qualification, those completing compulsory education, extended secondary education, undergraduate education or postgraduate education have a higher probability of being employed. In addition, both higher cognitive skills (as proxied by higher reading scores) and better family background (as proxied by father's higher education level) are associated with higher employment probability.

As stated above, the instrumental variable method is employed to address the possible issues of reverse causality and measurement error. I also test the exogeneity of the five personality traits, and the test results indicate that we should reject the exogeneity of the five personality traits (p=0.0000). The first-stage regression results are presented in Table 1.3. We can find that the instruments are jointly significant in all of the five first-stage regressions. Being bullied is negatively associated with extraversion. A child who is often

[21] Age is not controlled since the individuals from the NCDS are equally aged.

Ethnic group is not controlled because the sample is a highly homogeneous population, around 97.4% of the sample population is British. Also, I run robustness checks when the sample is restricted to British, and get very similar results.

Work experience is not controlled since this study is based on the first deposit of the NCDS data set, and the variables related to job history are not available when this chapter is completed. The variable work experience can be included in the model in future research. Also, the sample of this study is a very homogeneous population, all males, 97% British, same age, so we may expect that after controlling for education level, work experience may have little variation, and thus omitting work experience should not bias the estimation of the effects of personality traits.

[22] The changes in the employment probability associated with a one standard deviation increase in the independent variables for the probit specification are exhibited in the second column of Table 1.2.

bullied by other children usually cannot have good relations with others, and thus would tend to be an introvert as he grows up. Having difficulty in concentrating on anything for more than a few moments is positively correlated with extraversion. A child's immaturity is negatively correlated with agreeableness, since immaturity may suggest that the child is not considerate for others (including parents and teachers). Another instrument that is closely related to agreeableness is the dummy indicating whether the child is disobedient at home. It is hard to expect a child who is disobedient to be considerate or sympathetic towards other persons. When considering the instruments for conscientiousness, preferring to do things on one's own may help to enhance a child's independence and acceptance of responsibility, and consequently makes the child become more conscientious. However, being squirmy or fidgety, which is a signal of a lack in self-discipline and self-organization, is adversely associated with conscientiousness. All the instruments for emotional stability are adversely correlated with it. Children with hostility towards adults are usually very moody when asked by adults to do something, and a frequent mood swing is a typical reflection of unstable emotions. Worrying is an important component of emotional instability as well. In addition, being irritable exhibits that the person is not quite good at dealing with negative emotions. Lastly, when we consider the instrument for the imagination, biting nails, having nervous symptoms, or being upset by things happened for the first time may imply that it is difficult for the child to get used to new situation, and this may indicate that the person is not quite good at accepting new things and not quite creative or imaginative. An interesting finding is that an individual's educational attainment is positively associated with his personality scores: compared to those individuals with no academic qualifications, those completing secondary, extended secondary or tertiary education generally report higher scores on their personality scores. This finding suggests that an individual's self-reported personality measures may be affected by his education level.

The second-stage IVprobit regression results are shown in the last two columns of Table 1.2. Again, marginal effects are shown here. The null hypothesis of overidentification test of all instruments cannot be rejected, indicating the validity of the instruments. The results suggest an approximately 6 and 8 percentage points higher probability of being employed for an individual by a one standard deviation increase in agreeableness and conscientiousness respectively, but a 26 percentage points lower probability of being employed by a one standard deviation increase in imagination. Conscientiousness, a personality characteristic involving self-discipline, diligence and self-organization, is a key trait that improves job performance and productivity and thus greatly valued by employers in general. An agreeable person is usually considerate and cooperative towards other persons, and this is especially important in teamwork. Therefore, it is not surprising to find the positive effect of agreeableness on employment probability. The negative sign of the effect for imagination, nonetheless, may simply be a reflection of the unfavorableness of non-conformity and autonomy associated with imagination. An imaginative or creative individual who is full of his own ideas may not want to comply with existing rules and thus tend to be non-conforming (see arguments made by Johnson, 1983; Judge et al., 1999; Heineck and Anger, 2010). However, such non-conformity may not be welcomed by employers in many cases, leading to a penalty in the labor market.

When we compare the marginal effects in probit and IVprobit models, we find that conscientiousness and imagination exhibit greater effects on an individual's employment probability when estimated with instrumental variable technique, while emotional stability loses significance in the IVprobit specification. One thing worthy of notice here is that the sign of agreeableness changes from negative to positive from probit to IVprobit specification. This may serve as an evidence for the existence of reverse causality. The competition and stress in working may reduce one's consideration and sympathy towards other persons, and such reverse causality may even reverse the sign of agreeableness and let us observe a negative correlation between agreeableness and employment probability, while agreeableness actually imposes a positive effect on one's employment probability. For conscientiousness and imagination, the probit estimates are biased towards zero. This may be due to the existence of large measurement errors in the self-reported subjective personality traits.

Coefficients other than the personality traits have expected signs. Compared to those persons who do not have any academic qualifications, those completing extended secondary education or tertiary education are significantly more likely to be employed. Cognitive skills also affect employment probability significantly: a one standard deviation increase in the reading test score raises employment probability by 8 percentage points. As expected, a more advantaged family background as represented by father's higher education level would benefit an individual in a way of higher employment probability.

1.5.2 Personality and Occupational Choice

Having shown that certain personality traits do affect an individual's employment status, now let us move on to the next question: do personality traits affect an individual's occupation choice? It is natural to expect that individuals with various personality profiles sort into different occupations. For example, an outgoing individual who has good social interactions may choose to work in a managerial occupation, while an individual who is full of ideas may choose to work in a creativity-related field.

In order to analyze the effects of personality traits on an employee's occupational choice, a multinomial logit (MNL) model is estimated. The occupations are divided into four categories: managerial, non-manual, professional and manual categories. The Wald tests of combining alternatives are rejected, indicating no pair of the four occupational categories can be further aggregated. Moreover, Hausman and Small-Hsiao tests of the assumption of the independence of irrelevant alternatives (IIA) are passed, suggesting that the IIA assumption is not violated. The set of control variables is the same to that in the employment status model. The marginal effects for MNL are shown in Table 1.4. Extraversion imposes a significant impact on being in a managerial or manual group: a one standard deviation increase in extraversion is associated with a 5 percentage points increase in the probability of being in a managerial occupation and a 3 percentage points reduction in the probability of being in a manual occupation. Agreeableness is positively associated with the likelihood of being in a non-manual occupation but negatively associated with the likelihood of being in a managerial position. The more conscientious an individual is, the more likely he works in a managerial or professional occupation, but the less likely he works in a manual occupation.

However, these results only suggest the possible associations between one's personality traits and occupational choice, and the true relationship between personality traits and occupational choice may be biased due to the potential reverse causality and measurement error. Therefore, a two-stage multinomial logit (2SMNL) model is estimated. The joint significance of the instruments in all of the five first-stage regressions in Table 1.5 suggests the relevance of the instruments to the endogenous personality measures. However, the relative low F-statistics may suggest that the instruments are not correlated with the five endogenous personality measures quite strongly.

The second-stage results for the 2SMNL model are exhibited in Table 1.6. We find that three out of the five personality traits (extraversion, agreeableness and imagination) have significant effects on an individual's occupational choice.

Extraversion, the attribute associated with good social interactions, is appreciated in a managerial role. The more outgoing a person is, the more likely he takes on a leadership role. This effect is quite sizable: a one standard deviation rise in extraversion leads to a 30 percentage points increase in the probability of being a manager or employer. This is reasonable: a manager or employer needs to be good at organizing resources (including human resources), and such skill depends on good social interactions with other people (Ham, Junankar and Wells, 2009). Meanwhile, outgoing persons are less likely to be found in non-manual occupations: the probability of entering a non-manual occupation is reduced by 21 percentage points for a one standard deviation increase in extraversion. This may be due to the fact that the more outgoing a person is, the less likely he can settle down to work on his own, and for many non-manual workers, like artists, authors, photographers, etc., their main tasks are generally focused more on their own work than interacting with other persons.

Agreeable persons usually tend to be sympathetic toward others and concerned about pleasing others. However, this characteristic is not valued in professional occupations. A male worker who rates himself as one standard deviation more agreeable has an approximately 19 percentage points lower probability of working in a professional occupation. At first sight, this sounds surprising, since agreeableness seems to be a precious merit of kind-hearted persons. However, certain professional occupations do not favor individuals who are always concerned about pleasing others, like judges, police or military officers. Also, agreeableness and sympathy may be disadvantages for some other professionals, like accountants or actuaries who may be concerned more about competition than pleasing other persons (Ham, Junankar and Wells, 2009).

The last personality trait of importance, imagination, has a positive effect on being in a non-manual occupation and a negative effect on being in a managerial occupation, and the magnitudes of the effects are quite substantial: the probability of being in a non-manual occupation is raised by 29 percentage points, and the probability of being in a managerial occupation is reduced by 29 percentage points, for a one standard deviation increase in imagination. Imagination, a trait involving creativity, is an important characteristic for investigative or artistic activities (Barrick, Mount and Gupta, 2003), and thereby is greatly favored in many non-manual occupations which include lots of non-routine tasks. On the other hand, the non-conformity and autonomy associated with an imaginative personality may not be compatible with managing other workers.

Comparing the estimates of the five personality traits from MNL (Table 1.4) and 2SMNL (Table 1.6), we find that the significant personality traits in the 2SMNL specification have greater magnitudes than the corresponding MNL estimates. Meanwhile, some personality traits which are significant in the MNL specification lose significance when estimated by 2SMNL method.

Some other variables of interest have signs as expected. Education tends to have different effects depending on which occupational group is considered: a higher education level increases the probability of being in a non-manual or professional occupation, but reduces the probability of being in a manual occupation. Individuals with higher cognitive skills, as represented by higher reading test scores at age 11, are more likely to work as managers or employers, but less likely to work in manual occupations. Lastly, better family background, as represented by father's higher education level, leads to a higher probability of being in a professional occupation but a lower probability of being in a manual occupation.

1.5.3 Personality and Earnings

After establishing the effects of personality traits on employment status and occupational choice, now we explore how personality traits may help to explain earnings differentials among workers.[23] An OLS specification is employed first, followed by a 2SLS specification.

[23] In the main context, the dependent variable is the log of weekly earnings on the current main job. The

I also test the exogeneity of the five personality traits, and the results reject the null hypothesis that the five personality traits are exogenous (p-value=0.0077). I include the same set of control variables as in the employment status model.

The first two columns of Table 1.7 show the OLS regression results. Extraversion and conscientiousness are positively correlated with the male worker's earnings, while agreeableness is negatively associated with earnings. Both of the individual's own and his father's educational attainments are positively related to his earnings.

In order to address the possible endogeneity issue caused by reverse causality and measurement error, again, I employ the instrumental variable technique. The first-stage regression results are exhibited in Table 1.8. The F-tests imply the joint significance of the instruments in all of the five first-stage regressions, but the values of F-statistics are relatively low.

The last two columns of Table 1.7 present the second-stage regression results of the 2SLS specification for the earnings model. Extraversion, agreeableness, conscientiousness and emotional stability all have positive effects on earnings, while imagination has a negative effect on earnings. However, none of the Big Five personality traits is significant.

When focusing on the sample of full-time workers in estimating the effects of personality on earnings, we may need to worry about sample selection bias. For example, as demonstrated above, agreeableness, conscientiousness and imagination significantly affect an individual's probability of being employed, which would in turn affect his earnings. However, such indirect effects on earnings are not taken into account when we only consider the sample of workers and ignore those non-workers in the earnings model. Therefore, the effects reported in Table 1.7 (at least for agreeableness, conscientiousness and imagination) are understated and should serve as a lower bound of the true effects of personality traits on earnings.[24]

The effects of the other variables are in line with our expectations. Married workers get wage premiums in the labor market. Welsh workers earn lower earnings than those in

same earnings model is re-estimated when the dependent variable is the log of total weekly earnings, that is, the sum of earnings from the current main job as well as other jobs, and the results are very similar.

[24] When we take into account the indirect effects of personality traits on employment probability, agreeableness and conscientiousness should have positive effects on earnings, and imagination should have a negative effect on earnings, although the coefficients for these three personality traits are insignificant in Table 1.7.

England. Both higher cognitive skills and better family background yield wage premiums.

1.5.4 Extension

In the main context, the samples include both self-employed workers and paid employees. There may exist a concern that self-employed workers have different characteristics from paid employees, therefore, it would be informative to check whether the main findings about the effects of personality traits change if we exclude self-employed workers from the samples. I re-estimate the employment status, occupational choice and earnings models, with the samples excluding self-employed workers.

The results for the employment status model are shown in Table 1.9 and 1.10. When we compare Table 1.9 and Table 1.2, we find that our findings about the effects of the Big Five on an individual's employment probability are robust to whether including self-employed workers or not. The more agreeable and conscientious a person is, the more likely he is employed. The negative relationship between imagination and employment probability is also significant in Table 1.9.

Now let us check the effects of the Big Five on an individual's occupational choice when the sample excludes self-employed workers (Table 1.13). We find that most results are consistent to those in Table 1.6 when self-employed workers are included in the sample. The more extraverted a person is, the more likely he works in a managerial occupation but the less likely he works in a non-manual occupation. Imagination is positively associated with the likelihood of being in a non-manual occupation. Agreeableness is negatively correlated with one's probability of being in a professional occupation, but this effect loses significance in Table 1.13. One thing worthy of notice here is conscientiousness. The more conscientious a person is, the more likely he works in a managerial occupation, but the less likely he works in a non-manual occupation. These effects are consistent to those we find in Table 1.6, although they are not significant when we include self-employed workers. The negative relationship between conscientiousness and the likelihood of being in a non-manual occupation seems surprising, but this may be a reflection of the negative correlation between conscientiousness and intelligence. Intelligence is highly valued in non-manual occupations because these occupations involve many non-routine job tasks where intelligence is needed, therefore, conscientiousness is not valued in these occupations since it is negatively correlated with intelligence.[25]

For the earnings model, when the sample is restricted to paid employees, extraversion and conscientiousness impose positive influences on one's weekly earnings: if there was a one standard deviation increase in extraversion or conscientiousness score, the person's weekly earnings would rise by 41% or 36% respectively. Extraverts are generally more likely to take on leadership roles where the nature of being enthusiastic, outgoing and good at interacting with others is greatly appreciated (Ham, Junankar and Wells, 2009). This may impose a positive effect on their earnings. Conscientiousness is related to an individual's self-discipline, self-organization and hard working. Obviously this trait is welcomed by employers, and thus gets rewarded in the labor market. When comparing the 2SLS to the OLS estimates, it is clear that the OLS estimates for extraversion and conscientiousness are underestimated due to the combined effects of measurement error and reverse causality. Agreeableness and emotional stability, however, lose significance in the 2SLS specification.

When we compare Table 1.14 to Table 1.7, the coefficients for the Big Five personality traits have the same signs in both tables. However, while both extraversion and conscientiousness have significant effects for paid employee's earnings (Table 1.14), these effects are not significant when we also include self-employed workers in the sample (Table 1.7). A possible explanation is that personality traits have different effects on the earnings of self-employed workers and paid employees, and when we include both self-employed and paid employees in the sample, the effects of personality traits are messed up and lose significance.[26].

1.6 Discussion and Policy Implications

[25] Many studies have shown the negative correlation between conscientiousness and intelligence (see for example, Moutafi, Furnham and Crump, 2003; Moutafi, Furnham and Paltiel, 2004). As Moutafi, Furnham and Crump (2003) suggest, such negative relationship may be because less intelligent people cope with their weakness in intelligence by becoming more organized, diligent and responsible, while for those more intelligent people, they feel that they do not need to put in so much effort to improve conscientiousness, since they could rely on their intelligence to cope with most tasks in work and life.

[26] It would be informative to estimate the earnings model with the sample only including self-employed workers, and check whether the effects of the Big Five personality traits are different to those in the model when the sample only includes paid employees. However, this approach cannot be realized since the instruments are not valid when the sample only includes self-employed workers.

This study contributes to the literature by investigating the causal effects of the multi-dimensional Big Five personality traits on an individual's employment status, occupational choice and earnings, with data from the NCDS. In order to deal with the possible issues of reverse causality and measurement error, I employ instrumental variable approach, with the IVprobit specification for the employment status model, the two-stage multinomial logit specification for the occupational choice model, and the 2SLS specification for the earnings model.

One concern about the present study is that the instruments are not quite strongly correlated with the endogenous Big Five personality traits. Therefore, the IV estimates may be biased themselves, and they may be biased in the same direction of OLS estimates (Bound, Jaeger and Baker, 1995). Ideally, we would like to have some instruments that are closely correlated with adult personality traits, like the direct measures of childhood personality traits. However, the NCDS only provides the teacher's and mother's assessments for an individual's childhood behavior which can serve as proxies for childhood personality traits, but no direct measures of childhood personality traits. This is a limitation of the study.

The results indicate that personality traits play an important role in explaining the variation in labor market outcomes. The more agreeable, the more conscientious, and the less imaginative a person is, the more likely he is to be employed. The more outgoing and the less imaginative an individual is, the more likely he works in a managerial occupation, but the less likely he works in a non-manual occupation. Moreover, agreeableness reduces one's probability of being in a professional occupation. Considering the effects of personality traits on an individual's weekly earnings, both extraversion and conscientiousness lead to higher earnings for paid employees.

The results exhibit the importance of dealing with the endogeneity problem caused by reverse causality and measurement error. There exist great differences between the estimates when treating personality traits as exogenous and endogenous. For example, for the employment status model, the IV estimate for conscientiousness is about five times the probit estimate, and the IV estimate for agreeableness even has a different sign to the corresponding probit estimate. For the earnings model (for paid employees), 2SLS estimates for extraversion and conscientiousness are also much bigger than the corresponding OLS estimates.

My findings have important implications. The present study establishes that an individual's personality traits significantly affect his employment probability, occupational choice and earnings, even after controlling for the standard human capital and other background variables. While the traditional human capital variables and cognitive skills cannot account for much of the variation in labor market outcomes, this study shows the significant explanatory power of personality traits in one's labor market outcomes, which is of great theoretical importance. While this study can help us better understand the determinants of labor market outcomes, it may further contribute to our understanding on some important issues closely related to labor market outcomes, like wage inequality, poverty, and so on.

Practically, in order to help individuals gain more economic success, the education system and worker training programs may need to incorporate the training of behavioral and social skills, in addition to the training of cognitive skills.[27] In particular, when people are trained to interact well with others and be responsible and self-organized, they may be better equipped to deal with challenges in labor market. The Perry Preschool Program in the United States is a good example. It is a randomized experiment designed for disadvantaged young African American children, by teaching them skills of planning, organization and self-control. It is shown that the Perry Preschool Program helped shape the participants' personality traits, and in turn benefited the participants on a variety of later life outcomes (Heckman, Malofeeva, Pinto et al., 2010; Heckman, Moon, Pinto et al., 2010).

Also, the results in this chapter provide useful guidelines to job counselors on how to help workers with different personality traits find the occupations that are most suitable for them. Furthermore, workers now understand which personality traits are valued by employers, and could correspondingly adjust their behavioral and social skills to increase the probability of being employed and be better rewarded in the labor market.

[27] A similar implication is also mentioned by Nyhus and Pons (2005) and Osborne Groves (2005).

Chapter 2

Do Care Providers Sort into Flexible Jobs? Evidence from the Health and Retirement Study

2.1 Introduction

The rapid increase of the elderly population has raised great social concerns in the United States. A recent report by the United States Department of Health and Human Services points out that, due to the aging of baby boomers, the population over 65 years of age in the US has increased from 35 million to 40 million in the past decade, and is expected to further increase to 72.1 million by 2030. Given current population trends, by 2030 around one in five persons would be over 65 (U.S. Department of Health and Human Services, 2011). Such rapid expansion of the elderly population raises the question of who bears the heavy burden of elder care.

Traditionally, informal care provided by adult children and other family members is the common source of care to the elderly. This is the key means to "keep many individuals at home who would otherwise require expensive institutional care" (U.S. Department of Health and Human Services, 1997, P8). Adult children have assumed increasing importance as caregivers in recent years, due to the differential life expectancies of men and women and the resulting large number of widowed elderly women (Van Houtven, Coe and Skira, 2010).[1] However, most individuals providing care to their elderly parents are employed in the labor market at the same time. The competing time demands of work and care impose great challenges on working-age adult children. Studies have found that, care can significantly reduce paid employment by forcing care providers to either totally withdraw from the labor market or reduce work hours (See for example, Ettner, 1995, 1996; Bolin, Lindgren and Lundborg, 2008). Other studies have shown that employment can also negatively affect the likelihood of providing care (Boaz and Muller, 1992; Michaud, Heitmueller and Nazarov, 2010). Therefore, how to balance work and family is a major issue for care providers.

[1] According to the report by the United States Health and Human Services (2011), persons reaching age 65 have an average life expectancy of an additional 20.0 years for women and 17.3 years for men. 40% elderly women in 2010 were widows.

Flexible work arrangements are frequently proposed as an important means for accommodating the needs of working care providers (see for example, Heitmueller, 2007; Bolin, Lindgren and Lundborg, 2008). This raises an interesting question: what is the effect of elder care on an individual's job choice with respect to flexibility? In particular, compared to non-care providers, are care providers more likely to sort into jobs with flexible work arrangements? *Ex ante*, it is difficult to answer such a question, due to the possible two-fold consequences of flexible work arrangements. On the one hand, flexible work arrangements can obviously help reconcile work and care, and thus generate less pressured working conditions for care providers. On the other hand, flexible work arrangements are often associated with some negative consequences, such as lower wages, reduced promotion opportunities, and the like (Rhoads, 1993, P18; Heywood, Sieberty and Weiz, 2007). As a result, it remains an open question as to whether caregivers consciously sort into flexible jobs or not.[2]

This chapter fills the gap in the literature by examining the effect of care provision on an individual's seeking of flexible jobs. Based on the national representative longitudinal data from the US Health and Retirement Study (HRS), the present study analyzes whether the provision of care affects an individual's choice of a flexible job or occupation. However, there may exist an empirical challenge caused by unobserved individual heterogeneity. Certain individual time-invariant characteristics (e.g. personality) may affect both job choice and care provision. In this chapter, the time-invariant individual heterogeneity is controlled for by using fixed effects panel data model. The study finds that, for both men and women, care responsibility causes people to be significantly more likely to sort into flexible jobs or occupations, compared to non-caregivers. This result is robust to different measures of job flexibility and different care definitions. Women and men realize job flexibility through different channels. While women care providers are more likely to directly choose jobs with flexible schedules, men care providers are more likely to realize job flexibility indirectly by sorting into flexible occupation categories.

2.2 Background and Previous Findings

[2] In this chapter, "flexible job" refers to a job or occupation with access to flexibility.

Working-age adult children are the common sources for elder care. However, time is scarce and both working and caregiving are time-consuming activities. This leads to the competing time demands of work and care. On the one hand, researchers find that care provision can negatively affect one's labor market outcomes, including labor market participation, work hours and wage rates. Most of the literature examining the labor market consequences of informal care focuses on the extensive and intensive margin of labor supply. For example, Heitmueller (2007) uses data from the British Household Panel Study and finds that co-residential caregiving significantly reduces one's labor market participation by around 15%. He also finds that intensive caregiving (defined as providing at least 20 hours of care per week) imposes an even greater effect: a reduction in labor force participation of up to 26%. Using data from Household, Income, Labour Dynamics in Australia Survey, Bittman, Hill and Thomson (2007) point out care providers are more likely to reduce work hours or labor force participation due to care responsibility. Among Canadian participants of 1996 General Social Survey, Latif (2006) finds that care provision statistically significantly reduces females' work hours. For the US, Ettner (1995) employs the 1986-1988 US Survey of Income and Program Participation panel data and shows that co-residence with a disabled elderly parent has a substantial negative effect on female labor supply. She argues that this mainly reflects the withdrawal of care providers from the labor market. In another study, Ettner (1996) also finds a significant effect of non-coresidential care: non-coresidential parent care reduces female labor supply by around 12 hours each week. Some other studies using Asian data get similar results. For example, Do (2008) finds that among Korean laborers, intensive care (defined as providing at least 10 hours of care per week) reduces female caregivers' labor force participation. Using China Health and Nutrition Survey, Liu, Dong and Zheng (2010) find a very interesting result: a Chinese worker's employment probability and work hours would be significantly reduced due to caring for parents-in-law, but such effect is not significant when caring for parents. Other studies have shown that caregiving can reduce providers' wage rates. Among US care providers, Van Houtven, Coe and Skira (2010) find that providing care to elderly parents leads to a 3% reduction, or equivalently a loss of $0.37 for a woman's hourly wage. Bittman, Hill and Thomson (2007) document that care providers generally have low incomes than non-care providers. Heitmueller and Inglis (2007) and Do (2008) also find the existence of a wage penalty for caregiving in the UK and South Korea respectively. Considering the long-term consequence of caregiving, Wakabayashi and Donato (2006) shows that, compared to their non-caregiving counterparts, US caregivers are 25% more likely to live in poverty eight years after assisting elderly parents for personal care for at least 20 hours per week. Therefore, caregiving may impose substantial financial costs to care providers: they may lose income when they are forced to leave employment, reduce work hours or accept lower wages.

Compared to the literature analyzing the impacts of caregiving on labor market outcomes, fewer studies examine how one's employment affects his care provision. Time being scarce, the hours spent on paid employment will probably reduce the available hours for caregiving. Some studies suggest that employment can also reduce care provision. For example, among the participants of the National Long-term Care Survey and the National Informal Caregivers Survey, Boaz and Muller (1992) find that full-time employment reduces caregivers' unpaid help by 20 hours per week, but they do not find any effect of part-time employment on caregiving. Doty, Jackson and Crown (1998) also employ the National Long Term Care Survey and the National Informal Caregivers Survey, and show that paid employment significantly reduces care provision for females. For UK caregivers, Michaud, Heitmueller and Nazarov (2010) document a negative effect of employment on future co-residential and extra-residential caregiving decisions. Moreover, we may expect that one's wage level may have a negative effect on care provision, since the higher level of wage is, the higher opportunity cost of providing care is. Therefore, an individual with a high wage level may choose to purchase formal care services instead of providing care themselves. Couch, Daly and Wolf (1999) employ the 1988 wave of the Panel Study of Income Dynamics and find a negative relationship between wage rate and non-coresidential elder care for men and unmarried women.

Therefore, it is clear that care providers face a great difficulty balancing work and care. In order to help them integrate work and family responsibilities, many researchers have proposed flexible work arrangements (e.g., Heitmueller, 2007; Bolin, Lindgren and Lundborg, 2008; Berecki-Gisolf et al., 2008). Flexible work arrangements do not require workers to finish job tasks on a fixed schedule or at a fixed workplace. There are many types of flexible work arrangements. Typical flexible work arrangements include but are not limited to: flexibility in the scheduling of work hours, such as flextime and compressed work weeks; flexibility in the amount of work hours, such as job sharing and part-time work; and flexibility in the work place, such as working at home (Georgetown University Law Center, 2010). Many studies report that flexible work arrangements could help alleviate the burden of work-family balance, reduce stress, improve health, etc. (for example, see the arguments made by Hill et al. (2001) and Halpern (2005)). In particular, studies show that workers with access to flexible arrangements have lower rates of turnover and absenteeism, and are more likely to remain in the labor force (Kossek and Ozeki, 1999; Pavalko and Henderson, 2006). Flexible work arrangements can help to resolve the conflict between employment and care work for employed caregivers. However, flexible work arrangements may also be associated with negative consequences. The theory of compensating differentials suggests that workers may need to sacrifice earnings to gain job flexibility.[3] Using data from the US subset of the 1991 Comparative Project in Class Analysis, Heywood, Sieberty and Weiz (2007) find that flexible work arrangements are associated with approximately 20% lower earnings. In addition, flexible work arrangements may weaken employees' job-information networks and workplace interactions (Epstein et al., 1999), and consequently hinder their career advancement. Therefore, care providers need to consider the benefits as well as the costs associated with flexible work arrangements. Consequently, it is difficult to predict whether care providers would consciously self-select into flexible jobs, and this deserves an empirical examination.

2.3 Methodology

When examining the effect of elder care on an individual's job choice with respect to flexibility, we may meet an empirical challenge caused by unobserved individual heterogeneity. This is because certain time-invariant individual characteristics may be correlated with both care provision and job choice. For instance, an individual's distaste for pressure may

[3] From the employer perspective, they may offer lower wages in order to cover the costs (like administrative costs) associated with providing flexible work arrangements; from the employee perspective, they may would like to accept lower wages in exchange for the convenience brought by job flexibility to combine employment and care. For a more detailed discussion about the support and challenges faced by the theory of compensating wage differentials, see McCrate (2005).

make him sort into a job with flexible work arrangements, and at the same time, choose not to care for parents. Another possible component of individual heterogeneity that may be correlated with care provision is personality. For example, an emotionally unstable individual who is not good at dealing with stress and negative emotions may choose a flexible working environment and provide less care to generate a less pressured environment for himself. To account for such individual fixed effects, the following model with time-invariant individual heterogeneity as well as time fixed effects is estimated:

$$Flexible\ Job_{it} = \gamma Care_{it} + X_{it}'\beta + \alpha_i + \delta_t + u_{it}$$

where *Flexible Job* measures one's flexible job choice, *Care* represents the caregiving to parents and parents-in-law (if applicable).[4] The vector X includes a set of time-varying individual and household characteristics for individual i at time t. δ represents time-specific effects and is realized by including T-1 year dummies. α includes all time-invariant individual heterogeneity, and is allowed to be correlated with care provision. Once controlling for the time-invariant fixed effects, we need within variations in the care provision for each individual across time to identify the coefficient for the care variable.

2.4 Data and Measures

2.4.1 Primary Data

The primary data for this analysis are drawn from the eight panels (1996-2010) of the US Health and Retirement Study (HRS). The HRS is a panel study started in 1992. Personal interviews are conducted every 2 years with a representative sample of Americans over the age of 50. The HRS surveys five cohorts: the original HRS cohort born between 1931 and 1941, the Asset and Health Dynamics among the Oldest Old (AHEAD) cohort born in 1923 or before, the Children of the Depression Age (CODA) cohort born between 1924 and 1930, the War Baby (WB) cohort born between 1942 and 1947, and the Early Baby Boomer (EBB) cohort born between 1948 and 53. The survey includes cohort-eligible individuals

[4] I employ different measures of flexible job choice and different care definitions which will be described in detail in the next section.

and their spouses, if married, regardless of age. Different cohort members entered the HRS at different times.[5] The HRS provides rich information on an individual's employment and job characteristics, informal caregiving to elderly parents and parents-in-law, family structure, housing, income and wealth.

For the present analysis, I restrict the sample to those working-age individuals aged between 25 and 64 who are observed for at least two waves of the survey.[6] When the sample is restricted to workers, one might be concerned that excluding non-workers will lead to a sample selection bias. Being out of the labor force could be considered as the most flexible job choice. Studies have shown that care responsibility will cause some persons to withdraw from the labor market (see for example, Heitmueller, 2007; Bolin, Lindgren and Lundborg, 2008). Accordingly, when we focus on the working population, we actually underestimate the effect of care on flexible job choice. However, the persons who are in the labor market are those with the strongest attachment to the labor market, and this group is of our main interest. I exclude self-employed workers because one of the dependent variables (could adjust work hours in regular work schedule) is not available for self-employed workers in the original HRS data set.[7] The sample excludes observations from the 1992 and 1994 waves, because the definition of elder care in these two waves is inconsistent to that in the subsequent waves.[8]

2.4.2 Issues of Measurements

Dependent Variables. I employ two sets of dependent variables to measure job flexibility.

[5] The HRS cohort took the survey in 1992, 1994, and 1996, the AHEAD cohort took the survey in 1993 and 1995. Then from 1998, the HRS and AHEAD cohorts were merged and surveyed every other year. The CODA and WB cohorts entered the survey from 1998, and the EBB cohort entered the survey from 2004. More information about the HRS can be obtained from http://hrsonline.isr.umich.edu/.

[6] I also run robustness checks with the sample restricted to individuals in other age ranges, and get very similar results.

[7] Around 18% of the sample population aged between 25 and 64 who worked for at least two periods is self-employed.

I also run robustness checks for all model specifications when the sample includes both self-employed workers and paid employees, assuming self-employed workers as having flexible work schedule, i.e., the dependent variable "could adjust work hours" equaling 1. I get very similar results, indicating that excluding self-employed workers does not lead to sample selection bias.

[8] The 1992 wave does not include the information about the care of adult children to parents/in-law for their household chores, errands, transportation, etc., and such information are available for all of the subsequent waves. The 1994 wave asks respondents whether they provide at least 50 hours of care to parents/in-law in the past 12 months, while all the subsequent waves ask whether respondents provide at least 100 hours of care in the past 2 years.

The first dependent variable is a dummy variable from the HRS indicating an individual "could adjust work hours" which directly measures whether the respondent has work hour flexibility in the current job. The HRS respondents are asked about whether they have the kind of job where they could increase or reduce the number of paid hours in regular work schedule. They are considered as having hours flexibility if they could either increase or reduce work hours, and not having hours flexibility if they could neither increase nor reduce work hours.[9]

The second set of dependent variables includes some indirect measures of flexibility based on an individual's occupation category. One is the flexibility index for each occupation category, which is computed by the data from the May 1997 Supplement to the Current Population Survey (CPS).[10] The CPS is a monthly survey of around 50,000 house-

[9] Usually care providers need the ability to reduce work hours to balance work and care responsibilities. Therefore, "could reduce work hours" is a natural component of job flexibility. Here the ability to increase work hours is also considered as a component of job flexibility. This is because in some cases, when care providers have already reduced work hours to balance work and care, they may report as "could not reduce work hours" since they have already cut back work hours and cannot cut them further, but in fact they are in a job with access to flexibility. Therefore, if we only consider "could reduce work hours" as the measure for job flexibility, we may misunderstand the responses of these individuals. However, these individuals may report "could increase work hours" even if they have already cut back work hours due to care responsibility, and thus including "could increase work hours" as a component of job flexibility would provide a more precise measure of flexibility.

In addition, I also run regressions when the dependent variable defined only as "could reduce work hours", and get similar results to those presented in the main context.

[10] For the HRS data collections between 1992 and 2004, occupation was coded using the 1980 U. S. Census Occupation Code and masked for public release. There are 17 occupation categories, including managerial specialty operations, professional specialty operations and technical support, sales, clerical and administrative support, service occupations (private households occupations, cleaning and building service occupations), protection service, food preparation service, health service, personal service, farming and forestry and fishing, mechanics and repair, construction trade and extractors, precision production, machine operators, operators for transport and the like, operators for handlers and the like, and member of armed forces. For the HRS data collections between 2006 and 2010, occupation was coded using the 2000 U. S. Census Occupation Code and again masked for public release. There are 25 occupation categories in the HRS (Nolte and Servais, 2010). To keep consistency of analysis, the new 25 occupation categories are converted to the old 1980 occupation classification system. Some occupation categories in the HRS are aggregated for the ease of conversion. Specifically, "service occupations (private households occupations, cleaning and building service occupations)" and "personal service" are combined to a new occupation category as "personal care and service occupations", "precision production" and "machine operators" are combined to form a new occupation category as "production occupations", "operators for transport and the like" and "operators for handlers and the like" are combined to form a new category as "transportation and material moving occupations". The occupation category "member of armed forces" is omitted since it has a quite different nature to other occupation categories, members of armed forces are not the target working population that we are interested in. In addition, only 0.05% of the employed population is in this category, so it should not cause a bias when members of armed forces are dropped from the sample. After conversion, there are 13 occupation categories in the present study.

I also construct an occupation flexibility index using the data from the May 2001 Supplement to the CPS, since the May 2001 Supplement has similar questions about job flexibility to that in the May 1997 Supplement. I run robustness checks by estimating the main models with the flexibility index from 2001 CPS Supplement, and get quite similar results to those presented in the study.

holds conducted by the Bureau of Census for the Bureau of Labor Statistics. In the May 1997 Supplement, employed workers were asked about whether they have flexible work hours that allow them to vary the starting and ending times of their workday. Based on the responses to this question, I compute the weighted shares of workers with flexible schedules by occupation category, and treat these shares as indirect measures of flexibility occupation titles.[11] Based on this flexibility index, the occupation categories in the HRS are ordered to construct the last two dependent variables: "the top 3 flexible occupation categories" and "the least flexible occupation category". While there are 13 occupation categories in total in the study, the top 3 flexible occupation categories are those with the highest flexibility index which include managerial occupations, sales, professional specialty occupations and technical support, and the least flexible occupation category includes production occupations.[12]

Key Explanatory Variables. The HRS asks respondents about two types of caregiving to the elderly: personal activity assistance and chore assistance. Specifically, respondents are asked whether they or their spouses (if applicable) spent at least 100 hours in the past 2 years helping their parents or parents-in-law "with basic personal activities like dressing, eating and bathing". Respondents are also asked whether they or their spouses (if applicable) spent at least 100 hours in the past 2 years helping their parents or parents-in-law "with other things, such as household chores, errands, transportation, etc." For each type of assistance, a respondent who answers "yes" to the above questions is further asked about the amount of care he and his spouse (if applicable) each individually provided. Any respondent who cannot provide the precise number of care hours is asked to compare care hours with the 200 benchmark, i.e., whether he and his spouse's care hours are less than, equal to or more than 200 hours in the past 2 years.

The present study defines the total care hours for each individual as the sum of the time she spent helping the respondent's parents and parents-in-law (if applicable) for basic personal activities and household chores in the past 2 years. Based on the total care hours,

[11] Here, to calculate the shares, I use the individual weights provided by the CPS to make the sample population nationally representative.

[12] Here, professional specialty occupations and technical support include occupations like scientists, teachers, lawyers, technicians, etc. Production occupation category includes occupations like operators, fabricators, laborers, etc.

I define three sets of care variables: a dummy variable care100 (=1 if care hours$\geq$100); a dummy variable care200 (=1 if care hours$\geq$200); and a set of dummy variables care1, care2, care3 and care4, corresponding to the cases for care hours=0, 0<care hours<200, 200$\leq$care hours<500 and care hours$\geq$500 respectively. The purpose of employing different care definitions is to examine whether the final results are robust to different measures of care provision.

Other Explanatory Variables. In addition to the key care variables, I also control for some time-variant explanatory variables: age, age squared, health dummies (excellent health, very good health, good health, fair health and poor health, with excellent health as the reference group), region dummies (northeast, midwest, south, west and other regions, with northeast as the reference group), marital status (having a spouse versus not having a spouse), a dummy for having an employed spouse, a dummy for having a spouse with an Activities of Daily Living (ADL) limitation which is an indicator for spouse's health condition,[13] a dummy for having children younger than 18 years old in the household, household size, a dummy for home ownership, non-labor income in \$1,000,000, household non-housing wealth in \$1,000,000, work experience and experience squared.[14]

After eliminating missing data, the sample for one baseline model specification (with the dependent variable "could adjust work hours" and the key independent variable "care for at least 100 hours") includes 14,229 and 10,405 person-wave observations for women and men respectively. Depending on which measure of job flexibility and care provision are used, the sample size differs across model specifications.[15] This is shown at the main result tables 2.2-2.5.

[13] Here, Activities of Daily Living (ADL) includes bathing, dressing, eating, getting in/out of bed, and walking across a room. We may expect a person has very poor health if he had difficulties with ADL.

[14] Here, non-labor income is computed as the sum of capital income, pension, incomes from Social Security Disability, Social Security Retirement, other government transfers (including veteran's benefits, welfare, and food stamps), and some other incomes (e.g. alimony, inheritance, etc.). It is measured in \$1,000,000, i.e., non-labor income=2 means the individual's non-labor income is \$2,000,000.

Household non-housing wealth is computed as the sum of wealth components less debt. Specifically, it is the sum of the net values of real estate (excluding primary residence), vehicles, businesses, IRA (Individual Retirement Account), stocks, mutual funds and investment trusts, checking, savings and money market accounts, CD, government savings bonds and T-bills, bonds and bond funds, and other savings less debt. It is also measured in \$1,000,000.

[15] More details about the different measures of job flexibility and care provision are provided in the next section.

2.5 Empirical Results

2.5.1 Descriptive Analysis

Due to gender differences in attachment to the labor market, men and women are analyzed separately. Table 2.1a shows the summary statistics for the care variables by gender. Around 22% of the female sample provides care to the elderly for at least 100 hours, and around 15% cares for parents/in-law for more than 200 hours. Lower percentage of men provides elder care. 17% of the male sample cares for at least 100 hours, and 10% cares for more than 200 hours. Table 2.1b presents the summary statistics for the other variables of interest by gender and care status. The summary statistics are computed based on the baseline sample with the dependent variable as "could adjust work hours" and the care variable as "care hours$\geq$100". Compared to non-caregivers, female caregivers are more likely to be found in the top 3 flexible occupation categories (46.05%) than non-caregivers (41.72%), and are also more likely to be in a occupation category with a higher flexibility index and less likely to be in the least flexible occupation category, and these differences are statistically significant. For males, caregivers and non-caregivers differ little in the measures of job flexibility.

When comparing the individual and household characteristics for caregivers and non-caregivers, we get some statistically significant differences. Female caregivers are generally younger than non-caregivers. Caregivers have a 1.4 percentage point higher rates of marriage and a 1.4 percentage point higher rates of having an employed spouse. Males exhibit similar characteristics. The differences between caregiving and non-caregiving males are smaller than that for females, with one exception: 61% of male caregivers and 57% of non-caregivers have an employed spouse. Regarding the household configuration, compared to non-caregivers, male caregivers generally have smaller household sizes and are less likely to have kids younger than 18 years old, while female caregiver on average have larger household sizes. Care providers are in better health than non-providers. This pattern is more obvious for males: 56.36% of male care providers compared to 53.25% of non-care providers rate their own health as excellent or very good. Compared to non-caregivers, caregivers generally have higher rates of home ownership.

2.5.2 Main Findings

Due to the possible existence of unobserved individual heterogeneity, a fixed effects model is used to examine how elder care affects an individual's flexible job choice, with the results provided in Table 2.2 to Table 2.5.[16] When the dependent variable is a binary variable, a linear probability model with fixed effects is used. Although linear probability model has a well-known weakness as the predicted probabilities may lie outside the unit interval, I still use it in this study, because it is very easy to estimate and interpret marginal effects from a linear probability model when there exist individual fixed effects which are correlated with explanatory variables.[17]

Table 2.2 presents the results for the specification when the dependent variable is "could adjust work hours" in the current job. The first three and the last three columns correspond to the cases with different care definitions, for women and men respectively. From the first column, we can see that, compared to those women who do not care or provide a very

[16] All the analyses in the present study allow panel-robust standard errors that permit errors to be heteroscedastic as well as correlated over time for a given individual.

[17] It is also possible to estimate a logit FE model, but we cannot get marginal effects which are of our main interest, since fixed effects are never estimated.

I test whether a random effects model which imposes additional assumption that the time-invariant individual characteristics being uncorrelated with the explanatory variables is appropriate, and the test cannot be passed for most of the model specifications, indicating that a fixed-effects model is more appropriate in our context.

One more concern about analyzing the effect of elder care on flexible job choice is the possible existence of reverse causality, that is, an individual's being in a job with flexible work options may affect his decision to provide care. In order to test for such possibility, I try to use instrumental variable technique. The instruments include parental health condition and the numbers of sisters and brothers for the respondent. Here, parental health condition is proxied by whether parents/in-law need help with basic personal activities like dressing, eating, or bathing. These instruments are jointly significant in the first-stage regressions and pass the overidentification tests in the second-stage regressions for all model specifications. Based on the results from instrumental variable methods, we cannot reject the exogeneity of care provision for any model specification, that is, the possible reverse causality from flexible job choice to care provision should be negligible, if any. This is not surprising. Similar results are provided by many existing studies which worry about the endogeneity bias when analyzing the effect of informal caregiving on one's labor market outcomes. These studies also test for the endogeneity of care provision, and most of them cannot reject the null hypothesis of exogeneity (see for example, Heitmueller, 2007; Bolin, Lindgren and Lundborg, 2008; Van Houtven, Coe and Skira, 2010). However, the results in this study from instrumental variable methods are only suggestive since the instruments are not quite satisfying. Parental health condition is closely correlated with an individual's care decision, whether parents/in-law need care is a perfect proxy for the demand for care help from adult children. However, this instrument may not satisfy the exclusion restriction, since people who choose not to care for parents/in-law may be more likely to report that the elderly parents/in-law do not need care. For the numbers of sisters and brothers, they should not affect one's flexible job choice except through its effect on informal caregiving to the elderly, but they are not quite correlated with one's care provision. Therefore, we can only treat results from instrumental variable methods as suggestive and these results are not shown here.

small amount of care (fewer than 100 hours), care providers who help parents/in-law for at least 100 hours are 3.4 percentage points more likely to be found in a job with work hours flexibility. With a more detailed classification of caregiving based on the total care hours (Column 3), female caregivers with care hours ranging between 0 and 200 are 3 percentage points more likely than non-caregivers to sort into a job where they could adjust work hours when necessary. However, the effect of 0-200 care hours is not statistically different from the effect of 200-500. However, when we look at the last three columns, care provision does not significantly affect men's job choices with respect to flexibility. Therefore, the pattern emerges from Table 2.2 is that, while caregiving women are significantly more likely to choose a job with hours flexibility than non-caregiving women, such difference does not exist between male care providers and non-providers.

When using the indirect flexibility index by occupation category as the dependent variable (Table 2.3), caregiving imposes a significant effect on males' occupational choices: on average, compared to non-caregivers, male care providers select into occupation categories with higher flexibility index. However, no such difference exists for the female sample: regressors are jointly insignificant for the female model I, II and III.

Based on the flexibility index by occupation category, all the occupation categories are ordered and the last two dependent variables are generated as "the top 3 flexible occupation categories" and "the least flexible occupation category".[18] Tables 2.4 and 2.5 show the results when using these two variables as the dependent variable. From Table 2.4, we can find a significant though modest effect of caregiving on men's occupational choices: compared to non-caregivers or caregivers providing a small amount of care, male caregivers with relatively larger amount of care (care hours greater than 100) are about 2 percentage points more likely to choose the top 3 flexible occupation categories. From the last column, we can further find that, compared to non-care providers, male care providers who care for at least 500 hours are 3 percentage points more likely to sort into the top 3 flexible occupation categories. However, caregiving does not have any significant effect on women's occupational

[18] I also construct a dummy variable for "the bottom 3 flexible occupation categories", referring to the 3 occupation categories with the smallest flexibility index. However, regressors are jointly insignificant for the models with "the bottom 3 flexible occupation categories" as the dependent variable, for both men and women, and thus the results are not reported here.

choices. A similar story can be found in Table 2.5: male caregivers are significantly less likely to be found in the least flexible occupation category, but caregiving does not affect whether female caregivers sort into the least flexible occupation category.

Generally speaking, compared to non-caregivers, caregivers are more likely to sort into flexible jobs for both men and women. But they realize this goal through different channels: caregiving women are more likely to realize job flexibility directly by choosing a job where they could adjust work hours when necessary, while caregiving men are more likely to realize job flexibility indirectly by selecting into a flexible occupation category.[19]

The signs of other explanatory variables are in line with our expectations.[20] We note that household wealth is positively associated with the likelihood of being in a flexible job for females. As household wealth rises, women's ability to afford the possible negative consequences associated with flexible work options rises, and therefore they are more likely to choose a flexible job. In addition, women who are younger and have less experience in the job market are more likely to sort into flexible jobs. For males, those who have a spouse are less likely to select flexible jobs, probably because they can rely on their wife to take care of family responsibility and thus do not need to choose a flexible job to balance work and family responsibility, or because married men have greater financial obligations and therefore cannot afford a flexible job. Another pattern exhibited in the results is that non-labor income is positively associated with males' likelihood of being in a job with access to flexibility.

In order to check whether unobserved individual heterogeneity biases the OLS results of the relationship between elder care and job choice, I also estimate two baseline models with pooled OLS, with the dependent variables as "could adjust work hours" and "the top 3 flexible occupations" and the independent variable as "care hours$\geq$100", and then compare the results of OLS and FE. The results are shown in Table 2.6.[21] For simplicity, only the

[19] Since the regressors are jointly insignificant for the model Male III in Table 2.5, the results shown for this model specification is only suggestive.

[20] The indirect measures of job flexibility (flexibility index, the top 3 flexible occupation categories, and the least flexible occupation category) do not work quite well for the female models, regressors are even jointly insignificant for some specifications. Therefore, when analyzing the linkages between other explanatory variables and job flexibility for women, we will focus on the model specifications with the dependent variable as "could adjust work hours" in the current job (Table 2.2).

[21] In addition to all explanatory variables in the FE model, the OLS specification also include two more sets of time-invariant individual characteristics as explanatory variables: education and race.

coefficients for the care variable are shown in the table. We find great differences between the OLS and FE estimates. Let us look at the results from the "flexible schedule" model specification (the dependent variable is "could adjust work hours") first. In the fixed effects estimation female caregivers are 3.4 percentage points more likely than non-caregivers to find a job with flexible schedule. In the OLS estimation, the estimated coefficient is much smaller and not statistically different from zero. This difference may be because unobserved individual fixed effects (e.g., distaste for pressure) may be negatively correlated with care provision, and lead to a underestimation of the true effect of care provision. For males, there is no significant difference between the OLS and FE estimates for having flexible work hours. However, the coefficient for care provision is not precisely estimated for the male sample in either the OLS or the FE model. To see how individual fixed effects work for males, let us compare the OLS and FE results for the model where the dependent variable is "the top 3 flexible occupation categories". Here, a big difference exists: individual heterogeneity causes the OLS coefficient for care provision to have the opposite sign of the FE coefficient. I also compare the OLS and FE results for all other model specifications for both males and females, and in most specifications, there exist great differences between OLS and FE estimates. These results suggest that the OLS results are biased due to the existence of individual fixed effects, and thus it is more appropriate to use FE model to capture unobserved heterogeneity which may be correlated with one's care provision.

2.5.3 Robustness Analysis

In the main context, I restrict the sample to the individuals who are non-self-employed workers for at least two periods, and do not impose any restriction on individuals' parents/in-law. Therefore, those individuals with deceased parents and parents-in-law (if applicable) are included in the sample and treated as non-caregivers. However, we might worry that the individuals who have no care responsibility (due to deceased parents and parents-in-law) may have different job choices to those individuals who have care responsibility (due to alive parents/in-law) but choose not to care. To check whether this is the case, I re-estimate the fixed effects models with the sample further limited to the individuals who have at least one alive parent/in-law in the current wave and thus are at risk of caregiving. The results are presented in Table 2.7-2.10. Comparing Table 2.2-2.5 and Table 2.7-2.10, we get consistent results: elder care imposes a positive effect on one's flexible job choice, female caregivers choose jobs with hours flexibility while male caregivers sort into flexible occupation categories. Furthermore, the magnitudes of care effects in Table 2.7-2.10 are quite similar to those estimated with my original sample. For example, compared to non-caregivers, female caregivers are 4 percentage points more likely to choose a flexible job, male caregivers are about 1.5 percentage points more likely to be found in the top 3 flexible occupation categories.

2.6 Conclusion

This chapter has sought to estimate the effect of elder care on individuals' job choices with respect to flexibility. When examining how elder care affects one's job choice, we need to take into account the possible existence of unobserved individual fixed effects like personality which may be correlated with both one's job choice and care provision. With the longitudinal data from the Health and Retirement Study, the comparison between the pooled OLS and FE results provides the evidence for the existence of individual heterogeneity, and thus it is appropriate to use fixed effects model to control for time-invariant individual heterogeneity.

One concern that we have for this study is the potential reverse causality problem, that is, an individual's being in a job with flexible work options may affect his decision to provide care. If we think being in a flexible job would help an individual balance work and care and thus be able to provide more care, then our coefficient estimates for care variables will be upward biased. A natural way to deal with reverse causality is to use instrumental variable methods. Ideally, we would like to have some instruments closely correlated with care provision but uncorrelated with the error term, like whether an individual observed her parents take care of her grandparents/in-law when she was a kid. However, the HRS does not provide satisfying instruments for care provision. Therefore, it is worthwhile to deal with reverse causality in future research.

Another legitimate concern is about the expansion of Medicaid-funded home- and community-based care services. In recent years, many states are expanding Medicaid to home- and community-based services for long-term care (Engquist et al., 2010). While more home- and community-based options paid by Medicaid are available for long-term care, adult children may choose to provide less care themselves. At the same time, Medicaid-funded care will release adult children from the financial responsibility of purchasing formal health care for their elderly parents, which may allow them to sort into flexible jobs more freely to generate a less pressured environment for themselves. If this is the case, then when the omitted expansion of home- and community-based care services is captured by the error term, our estimates of the effects of elder care provision to an individual's flexible job choice will be underestimated.

The study finds that for both men and women, elder care has a significant positive effect on choosing job flexibility. That is, compared to non-caregivers, caregivers are significantly more likely to sort into jobs or occupations with flexible work arrangements. This result is robust to different measures of job flexibility and different care definitions. Women and men realize job flexibility through different channels. While women care providers are more likely to directly choose jobs with flexible schedules, men care providers are more likely to realize job flexibility indirectly through sorting into flexible occupation categories. This implies that for caregivers, the benefit of integrating work and family responsibility outweighs the possible negative consequences associated with flexible work arrangements. Therefore, workplaces may want to provide more flexible work arrangements to help caregivers better balance paid employment and unpaid care work.

Chapter 3

Long-term Health and Socioeconomic Consequences of Child Labor: Evidence from Brazil

3.1 Introduction

Child labor is one of the most controversial issues in recent years. According to International Labor Organization (ILO)'s estimation (International Labour Organisation, 2010), in 2008, there were about 215 million children aged 5 to 17 years old working worldwide, with approximately 115 million engaged in hazardous work. Such alarming figures have attracted substantial attention, leading to calls for actions to be taken to deal with child labor.

Although sometimes child labor is an ethical issue seeming beyond discussion, identifying the long-run health and socioeconomic consequences of child labor is essential, since early entry into the labor force may affect an individual's income, health and education in his adulthood. However, much of the research on the consequences of early working emphasizes the short-run effects (see for example, Graitcer and Lerer, 2000; Milcent, Huguenin and Carusi-Machado, 2005; Ray and Lancaster, 2005), and the interactions between child work and adult health and socioeconomic status have not been widely explored. This is due to the limited data linking child work experience and adult outcomes.

The present study analyzes the interactions between participation into the labor market during childhood and an adult's income, health and educational attainment in Brazil. Data from Brazil Living Standards Measurement Study Survey (*Pesquisa Sobre Padrões De Vida*, PPV—1996/97) are employed. The key feature of this survey is that it asked each respondent at which age he started to work for the first time. This enables me to correlate early working to an adult's current health and socioeconomic conditions. However, instead of employing a dummy variable indicating whether the respondent was ever a child laborer, I include the age at which the individual started the first job in the model. This is because there is no agreement upon the definition of child labor, i.e., under which age we define a worker as a child laborer, and the results are sensitive to the definition of child labor if I include the dummy indicating whether the person ever worked in childhood. By including the variable age started to work, I can explore the long-run effects of one year earlier of entry into the labor market, and thus provide implications about the effects of child labor.

I study the impacts of child labor on three dimensions of an adult's status: income, health and schooling. To the best of my knowledge, few studies have ever looked at the multidimensional long-term consequences of child labor before. Researchers explore either just the linkages between early working and adult income (see for example, Ilahi, Orazem and Sedlacek, 2001; Emerson and Souza, 2007), or the relationships between child labor and adult health (e.g., Kassouf, Mckee and Mossialos, 2001; Lee and Orazem, 2010). Beegle, Dehejia, and Gatti (2005) examine the income, health and schooling effects of child labor in Vietnam, but their findings are limited to examining the outcomes only 5 years after child working and thus cannot provide the consequences of early working in the long-term view. Few existing literature provide a full picture of the long-run impacts of child labor, because most of the analyses to date pertain to different samples, different data sets and even different countries. However, this chapter complements the existing literature by studying the long-term income, health and schooling effects of child labor simultaneously with the same data set from Brazil. This enables the analysis to be done in the same framework and provides a coherent story about the possible long-term consequences if one enters the labor market early in her childhood. This is useful for policy discussion. When policy makers determine whether and the extent to which we should reduce the incidence of child labor in Brazil, this study provides a good reference about the multidimensional long-term effects of early working. In addition, I try to deal with the possible endogeneity problem by applying the instrumental variable method. Furthermore, as will be shown below, while most of the previous studies analyzing child labor in Brazil pool the urban and rural samples, I find great distinctions of child labor effects on urban and rural residents, which suggests the care needed for implementing child labor policies in different areas.

The present study finds that one year later entry into the labor market is associated with higher incomes, better self-assessed health indexes for rural adults and lower probabilities of getting health problems for both urban and rural adults. As for the schooling effect, the later one enters the labor market, the more years of schooling he obtains. While both the income and health impacts of child labor on urban residents are smaller than those on rural residents, urban residents suffer greater adverse schooling impacts than their rural peers. The main findings for the health and schooling impacts of early working are consistent when either the working sample (including workers with valid income information) or the full sample (including all individuals with valid, missing or zero income data) is employed for estimation.

3.2 Child Labor in Literature

My research is built upon a growing literature about the short-term and long-term consequences of child labor.

3.2.1 Child Labor and Schooling

Most of the current literature on child labor and schooling focuses on the relationships between early working and the contemporaneous schooling attendance and educational achievement.

Some evidence suggests that early entry into the labor force is negatively correlated to school attendance. Psacharopoulos (1997) observes that child labor makes working children receive 2 fewer years of schooling than their non-working peers in Venezuela. Based on data from Ghana in the late 1980's, Boozer and Suri (2001) conclude that there is a significant trade-off between working and attending school: one more hour of child work is associated with 0.38 fewer hour of school attendance. Assaad, Levison and Zibani (2001) also find a strong association between early working and school dropping out in Egypt.

The weight of evidence suggests that cognitive attainment from schooling is lower for working children, probably because working takes up part of the children's time and leaves children tired and less able to study effectively. Heady (2003) explores the linkages between early working and children's learning achievement with GLSS2 data set from Ghana. He suggests that working outside the household adversely affects children's results on reading and mathematics tests. Gunnarsson, Orazem and Sánchez (2006) estimate that child labor reduces math and language scores by 7.5% and 7% respectively, on 3th and 4rd graders in 11 Latin American Countries.

3.2.2 Child Labor and Adult Health

Many health risks caused by early working need time to manifest themselves. For example, the stress or negative emotions facing young laborers today may not have an immediate impact, but lead to depression or other psychological problems in their later life. On the other hand, however, the long-run health consequences of early working can be positive as well, because income from child work may be crucial to an extremely poor household (Psacharopoulos, 1997), and children's income contributions to the family may improve their living standards and nutritional status, and hence impose a positive impact on their long-run health development (for a more detailed discussion, see O'Donnell, Rosati and Doorslaer (2005)).

Previous research has examined the long-term health consequences of child labor. Kassouf, Mckee and Mossialos (2001) using a Brazilian data set find that as one enters the labor market earlier, his likelihood of reporting less than good health in adulthood increases. Based on an analysis of the Brazil PNAD data set, Lee and Orazem (2010) argue that early entry into labor market and decreasing schooling time jointly increase the probability of reporting physical ailments in adulthood.

3.2.3 Child Labor and Adult Income

Early exposure to work may affect a child worker's future income through human capital investment. Education provides skills that raise an individual's productivity and in turn raise his earnings. Therefore, how early working affects a child's education will have a link with his future income. Also, if child work leads to physical injury or psychological stress which may survive through adulthood, or if there is any health benefit arising from the improved nutritional status or living standard owing to young worker's income, such a health effect will affect future earnings in adulthood. In addition, when a child works early in his life, he is able to accumulate working experience which may have pecuniary benefits (for a more detailed discussion, see Emerson and Souza, 2007).

The linkages between child labor and subsequent labor market outcomes have been examined empirically but still many questions remain. Ilahi, Orazem and Sedlacek (2001) explore a national survey in Brazil (PNAD) and find that early exposure to child labor significantly reduces adult earnings and gives rise to an increase in the probability of being in poverty. However, their study does not take into account the possible endogeneity problem: there may exist some unobservable factors like an individual's ability that affect both the child labor decision and income. Hence, Ilahi, Orazem and Sedlacek's work only provides suggestive results. Also based upon the PNAD data set from Brazil, Emerson and Souza (2007) examine whether child labor imposes negative effects on adult earnings. They employ the GMM IV method to address possible endogeneity and find that child labor significantly reduce adult earnings for males even after controlling for schooling. However, Emerson and Souza do not control for the individual's health which may affect her income, and it would be of interest to consider whether early working could affect an adult's income when controlling for both schooling and health status, i.e., whether child labor has an income impact other than through its impacts on education and health. In addition, Emerson and Souza only focus on the income impact and omit other impacts of child labor. As will be shown below, my work complements their study by analyzing multidimensional consequences of early entry into the labor market.

3.2.4 Child Labor in Brazil

There is a long tradition of child labor in Brazil. The first registered child labor dates back to the 16th century, when children helped adults extract pau-brasil (the native Brazilian tree) (Ferreira, 2001). Along with the industrialization in the 20th century, there existed a great demand for child labor, and child employment became very serious through the whole century. According to Moura (1982), in 1912, 30% of the labor force in the four major textile factories was made up of children and adolescents, and this proportion even increased to 40% by 1919.

Although a sharp decrease of child labor occurred in the second half of 1990s, owing to the government's efforts to reduce child labor (such as "Bolsa Escola" which is a cash transfer program conditional on school attendance), there are still a large number of children involved in working. According to the estimation of Instituto Brasileiro de Geografia e Estatística (IBGE, 2007), there are about 5.4 million children aged between 5 and 17 years old in the labor market, of whom 40.7% are under 14 years old despite of the prohibition of child laborer younger than 14 years old from Federal Constitution of Brazil. Among those working children between 5 and 17 years old, one third work 40 hours or more per week. To be more specific, 13.6% of the 10-14 age group and almost one half of the 15-17 age group work more than 40 hours per week (ICFTU, 2004).

3.3 Data and Descriptive Analysis

The main data used for analysis come from the Living Standards Measurement Study Survey (*Pesquisa Sobre Padrões De Vida*, PPV—1996/97) of Brazil. The PPV was undertaken by the Instituto Brasileiro de Geografia e Estatística (IBGE) and the World Bank jointly from March 1996 to March 1997.

The PPV covered information from urban and rural areas in Northeast and Southeast of Brazil. The living standards in the Northeast Region are the lowest, while the Southeast is the richest region in Brazil. Hence the PPV provided two typical regions with respect to living standards and employment in Brazil. The survey interviewed 4940 households in total, collecting detailed information on household composition, migration, education, health, economic activity, fertility, etc. The key feature of the data set is that it asked each respondent about the age at which he started working and the working sector of his first job.[1] The PPV provided detailed information about an individual's health and socioeconomic status as well as his first job, meeting the requirement for my research question. It is noteworthy that the PPV has been little explored in this direction. While most studies on child labor in Brazil employ the Pesquisa Nacional por Amostra de Domicilios (PNAD) (e.g., Ilahi, Orazem and Sedlacek, 2001; Emerson and Souza, 2007; Lee and Orazem, 2010) data set, my analysis using the PPV could supplement the literature.

However, one potential weakness is that the information about a person's first job comes

[1] A person who has worked previously is understood as someone who:
(1) has exercised an economic activity paid in money, merchandise, products or only in benefits (housing, food, clothing, etc.);
(2) has exercised an economic activity with no payment for at least 1 hour per week for the purpose of helping a member of the household unit who has an economic activity, or as an apprentice, trainee, etc. (IBGE, DPE, and DEPIS, 1997).
The main working sectors include: agriculture, services, manufacturing, construction, textile, transportation and some other industries.

from a recall question and may be subjected to recall errors. Ideally, I would like to have a longitudinal survey in which the same persons are followed from their childhood to adulthood, as well as more detailed information about their first jobs are interviewed, like the working hours, working conditions, working and schooling, etc. But such data are rare, especially in developing economies where child labor is prevalent.

The sample is composed of individuals aged between 18 and 55 years old with valid information on earnings, health and education. I restrict the analysis to individuals older than 18 years since I want to analyze the impacts of early working on adults, and also to individuals younger than 55 years old, since 55 is the retirement age in Brazil and most people older than 55 do not have regular monthly earnings. Furthermore, my sample is selected to include individuals who entered the labor market between 5 and 31 years old, those persons who started to work younger than 5 years old or older than 31 years old are treated as outliers and thus dropped.[2] The sample size is 3901 after this selection process.[3] Owing to the distinct differences in urban and rural areas, all analyses are conducted separately for these two areas, with 3235 and 666 individuals respectively.

The summary information for the variables is presented in Table 3.1. Urban people typically have higher monthly earnings than rural residents.[4] As for the self-assessed health index, it equals 1 if the individual rates his own health condition as "poor" or "average", equals 2 if the individual rates health as "good" and equals 3 if the individual rates health as "very good" or "excellent".[5] From the summary statistics, urban and rural residents report very close and high health index: about 2.3, implying people evaluate their health conditions as more than good on average. In both urban and rural areas about two out of ten adults report to have health problems. Noticeably, there is a big gap of education levels between the urban and rural sample. While the average years of schooling in the urban sample is 8.33 years, implying people in urban areas complete upper primary education,

[2] Restricting the range of age started to work does not greatly reduce the number of observations compared to the original survey data, since only less than 1.2% of persons started to work before age 7 or after age 31. Also, my results are robust to modest changes in the range of age started to work.

[3] The sample size is also reduced due to the unavailability of data for instrumental variables in some years.

[4] While rural residents earn about 281 Reais each month on average, the average income of urban residents is 658 Reais, more than twice of that of rural adults.

[5] The categories "poor" and "average" are combined because few individuals report poor health. The categories "very good" and "excellent" are combined because few rural individuals report excellent health.

rural adults receive less than 5 years of education.[6] On average, urban residents entered the labor market at 15 years old, almost 3 years later than rural residents did. The composition of race for the urban and rural samples are quite similar. As for parental education levels, individuals in rural areas typically have less educated parents than those in urban areas.

Figure 3.1 presents the distribution of age started to work for the urban and rural samples. Note that rural individuals typically entered the labor market earlier than urban individuals. Figure 3.2 and 3.3 show the average of log-earnings and the average of the self-assessed health indexes by the age of labor market entry for urban and rural individuals respectively. In these figures, I collapse the individuals who started to work before 7 years old into one group and after 20 years old into another group, since from Figure 3.1, it is noteworthy that there are very few individuals starting their first jobs before 7 or after 20 years old, 91% of my sample entered the labor market between 7 and 20 years old. Figure 3.2 and 3.3 exhibit roughly linear relationships between log-earnings and starting age, and between the health index and starting age respectively.

The averages of years of schooling by age started to work are presented in Figure 3.4a. Again, the individuals who started working before 7 or after 20 years old are collapsed into two groups separately. The increase in the years of schooling associated with the increasing starting age is notable. However, this trend becomes complicated when I consider different quantiles of years of schooling. From Figure 3.4b, it is clear that the distributions of years of schooling differ conditional on different starting ages, for both urban and rural residents. This suggests that quantile regression is necessary for analysis.

3.4 Methodology

3.4.1 Child Labor and Adult Income

The model to be used for analyzing the long-run effect of child labor on adult income

[6] Brazilian education system: primary education (1st grau) consists of 1st-8th grade; high school education (2nd grau) consists of 9th-11th grade; undergraduate education typically consists of 4 years of schooling; graduate education differs according to degrees and fields. In this chapter, I split the primary education into lower primary (1st-4th grade) and upper primary (5th-8th grade) education as Emerson and Souza (2007) do.

is:

$$\ln \text{minc}_i = \pi_1 + \alpha_1 * startage_i + \beta_1 * sch_i + \gamma_{1a} * ehealth_i + \gamma_{1b} * ghealth_i + x_{1i}^{'}\delta_1 + \varepsilon_{1i} \tag{3.1}$$

where lnminc is the log of monthly income, startage is the age at which the person started the first job, sch is years of schooling, ehealth and ghealth are the dummies for reporting "excellent health" and "good health" respectively, and x_1 is a vector of exogenous variables, including age, age-squared, gender, race, GDP per capita of the individual's residence state at the interviewing year, and parental education levels which are a proxy of the individual's family background.[7]

It is likely that a person's decision to work, years of schooling and current health status are correlated to the unobserved components of income in model (3.1). For instance, an individual with higher ability tends to achieve higher level of education and earn higher income (biasing β_1 upward); an individual's unobserved health endowment is not only correlated to her current health status, but also affects her earnings and in turn biases the health coefficient. As for the coefficient on the age started to work, higher ability may lead to later entry into the labor market for an individual since she has the capacity to acquire higher levels of schooling (biasing α_1 upward). A higher unobserved health endowment, nonetheless, makes the child more likely to be sent into the labor market early and thus biases α_1 downward. Meanwhile, measurement error may also cause potential bias, making the directions of biases for the coefficients unpredictable. My data on people's age started to work come from a recall question, and thus would probably be subjected to recall bias. The possible measurement error in the self-reported health status will be discussed in the next section.

One way to address the possible endogeneity and measurement error is to employ instrumental variable technique. Variables qualified to be instruments must be sufficiently correlated to people's child labor and schooling decisions as well as current health conditions, but not correlated to the unexplained components of income.

[7] I include the linear specification of age started to work because the relationship between log-earnings and age started to work is roughly linear from the raw data. I also tried the model including starting age and squared starting age, and got qualitatively similar results.

One possible set of instruments for the age of labor market entry and years of schooling include parental occupations, the availability and quality of local education systems and the economic conditions in local labor markets when the individual was a child. Parikh and Sadoulet (2005) argue that children of employers or self-employed persons are more likely to work than children of employees. Thus parental occupations could affect an individual's decision to work in his childhood. The weight of evidence suggests that school quality is an important determinant of an individual's schooling decision (Bedi and Edwards, 2002), and the conditions in local labor markets will affect the supply of and demand for child laborers directly. Therefore, the instruments I use for age started to work and schooling are the parental occupations when the individual was 15 years old (denoted by z_1), the number of teachers per school for the state where the individual lived when he was 7 and 11 years old, since age 7 and 11 are the typical ages for a child to enter the lower primary and upper primary education in Brazil, and the GDP per capita of the state where the individual lived when he was 12 years old (denoted by z_2), since age 12 is the minimum age at which a child could legally enter the labor market in Brazil[8]. The choice of instruments about the state-

[8] Here, "typical" means the individual enters school at 7 years old and there is no delaying or repeating of grades.

85.41% of my sample consist of individuals whose current state of residence is the birth state. I assume that these individuals were not migrants and thus the birth state's figures of teachers and GDP per capita are used as instruments.

For an individual whose current residence state is not the birth state, I can identify the last state he lived before moving to the current state. If the last state the individual lived before he moved to current state was the birth state, then I assume that he just migrated from the birth state to the current state directly and migrated once in total; if the last state the individual lived was not the birth state, then he migrated at least twice in total.

I can identify how long an individual lived in the birth state. If the individual lived in the birth state for longer than 11 years, then no matter how many times he migrated, all the instruments use the birth state's information.

For those people who migrated from the birth state to the state of current residence directly, if he lived in the birth state for 11 years, then the state GDP per capita when the individual was 12 years old employs the figure from the current state and all the other instruments use figures from the birth state; if the individual lived in the birth state for longer than 6 years but shorter than 11 years, then the number of teachers per school when the individual was 11 years old and the state GDP per capita when the individual was 12 years old use figures from the current state while the rest instruments use figures of the birth state; if the individual lived in the birth state for shorter than 7 years, then all the instruments employ figures of the state of current residence.

For those people who migrated at least twice and lived in the birth states for shorter than 12 years, I cannot determine in which states they lived before coming to the current state and when they came to the current state, so information from the birth state are used as instruments for simplicity. However, there are only 33 individuals (0.85% of the whole sample) migrating at least twice and living in the birth states for shorter than 12 years, such a small portion of the sample should not affect my main results. I run a robustness check in which I give these persons current states' information as instruments, the results are similar.

level schooling and labor market conditions is guided by Emerson and Souza (2007) and Lee and Orazem (2010)[9]. But they all apply the birth state's information for instruments, while in the present study, I take advantage of both the birth state's and the current living state's data to construct the instruments which makes my instruments more informative. It is a challenge to find instruments which are exogenous to the unexplained components of income.[10] Once controlling for family background, current labor market conditions as well as other covariates, parental occupations, the variations of schooling quality and local labor market conditions when the individual was a child should be uncorrelated to the error term and satisfy the exclusion restriction.

Considering the instruments for health condition, I use the availability and quality of local health systems when the individual was a child which are represented by the numbers of hospitals, beds and doctors per 1000 inhabitants of the state where the individual lived when she was 7 years old (denoted by z_3). Controlling for all regressors, including the family background and current local labor market conditions, the availability and quality of local health systems when the individual was a child should not have independent influence on adult earnings. Furthermore, as I will demonstrate below, the relevance of the instruments for child labor and schooling decisions as well as the health conditions are checked through the tests of excluded instruments in the first-stage regressions, and the validity of instruments are checked through overidentification tests in the second stage.

Here, data on the parental occupations come from the PPV survey directly. Data on the number of schools and teachers, the number of hospitals, beds and doctors by state and year come from the IBGE online resource "Statistics of the 20th Century".[11] Data on the

[9] For the instruments of child labor and schooling decisions, Emerson and Souza (2007) employ the number of schools per 1000 children and the number of teachers per school in the birth state when the individual was 7, 11 and 15 years old, and the birth state's GDP per capita when the individual was 12 years old; and Lee and Orazem (2010) employ the number of schools per 1000 children and the number of teachers per 1000 children in the birth state when the individual was 7 years old, and the state-specific average wage rates for low-skilled workers in the year when the individual was 12 years old.

[10] Some people may argue that there may exist a persistency of occupation across generations and this causes parental occupations to be correlated to the unexplained components of income. However, I think that such intergenerational persistency of occupation mainly comes from the effect of parental education upon kids, and after controlling for parental education and other covariates in the model, parental occupations should satisfy the exclusion restriction.

[11] These series are available on line at http://www.ibge.gov.br/seculoxx/default.shtm (accessed on 09/11/2010).

GDP and population by state and year are taken from the IPEA historical series.[12]

To estimate the income model, I first run OLS regressions, and then employ 2SLS technique to rule out the possible endogeneity and measurement error. Comparing coefficients from OLS and 2SLS tells us the direction and magnitude of bias in the coefficient estimates due to the endogeneity problem and measurement error, if any.[13]

3.4.2 Child Labor and Adult Health

I employ two health models with two health indicators: one is the self-assessed health index and the other is the incidence of health problems. Self-reported health status has been shown to be a good proxy for a person's true health condition. Kalpan and Camacho (1983) and Mcgee et al. (1999) find persistent associations between self-reported health ratings (like poor, fair, good, excellent, etc.) and mortality, and self-reported health status is a strong prognostic indicator for subsequent mortality. Miilunpalo et al. (1997) reinforce this view and further show that the perceived health is inversely associated with the number of physician contacts per year. However, measurement error may exist in the studies employing self-reported health measures, since how people evaluate his health may depend on his education level, working status, etc. An individual with a higher education level is more likely to take care of himself and may have more information on his health condition. By examining the relationship between a self-reported health measure and a simulated clinical measure with the tetrachoric correlation coefficient, Butler et al. (1987) find the existence of biased reporting. In particular, non-working persons are more likely to report incorrect health conditions, probably due to the need of justification of unemployment.

The purpose for employing two health indicators is to mitigate the possible measurement error in the self-assessed health index since the incidence of health problems is relatively more accurate and objective than the health index. However, the incidence of health problems may not reflect the overall health condition as the health index does. Furthermore, I could check the consistency of results across models with two health measures, i.e., whether I can get a coherent story about the effect of child labor on adult health from different

[12] These series are available on line at http://www.ipeadata.gov.br/ipeaweb.dll/ipeadata?65370046 (accessed on 09/11/2010).

[13] All models in this chapter allow for clustering on the birth year and state.

health measures.[14]

The first health measure is an individual's self-reported health index. This ordered measure comes from the individual's self-assessed health status, that is, respectively, poor or average (*chealth*=1), good (*chealth*=2) and very good or excellent (*chealth*=3). I use the following ordered probit model to capture how the child labor decision affects an individual's health in her adulthood:

$$chealth = \begin{cases} 1 & \text{if health*} \leq \zeta_1 \\ 2 & \zeta_1 < \text{health*} \leq \zeta_2 \\ 3 & \text{health*} > \zeta_2 \end{cases}$$

$$health_i^* = \alpha_{2a} * startage_i + \beta_{2a} * sch_i + x_{1i}^{'}\delta_{2a} + z_{3i}^{'}\eta_{2a} + \varepsilon_{2ai} \qquad (3.2a)$$

When the latent health status variable *health** crosses a cutoff point, the observed category of the health index changes. x_1 and z_3 consist of the same variables as in the income model (3.1).

The incidence of health problems is used as the second health measure. Formally, the health problems reported in this survey include flu/cold/pneumonia, infection, accident/injury, digestive problem, pain, infarction and some other problems. Child labor may adversely (or positively) affect the incidence of health problems by affecting a young laborer's health capital and making him more (or less) likely to get a health problem in adulthood. A probit model is employed to estimate the effect of child labor on the incidence of health problems:

$$hproblem = \begin{cases} 1 & \text{if hproblem*} > 0 \\ 0 & \text{otherwise} \end{cases}$$

$$hproblem_i^* = \pi_{2b} + \alpha_{2b} * startage_i + \beta_{2b} * sch_i + x_{1i}^{'}\delta_{2b} + z_{3i}^{'}\eta_{2b} + \varepsilon_{2bi} \qquad (3.2b)$$

[14] Lee and Orazem (2010) also use multiple health measures to mitigate measurement error and check consistency of results across different health measures.

where *hproblem* implies the incidence of health problems and $hproblem^*$ is the latent variable. x_1 and z_3 are the same sets of variables as in the income model (3.1).

However, I am still faced with an endogeneity problem caused by unobservable health endowments. Only healthy children are qualified for employment which induces a positive relationship between health endowment and early working. Additionally, measurement error still remains a problem and makes the direction of bias on the coefficient estimates unpredictable. As a result, an IV ordered probit specification for model (3.2a) and an IV probit specification for model (3.2b) are employed to rule out potential bias and investigate the true health effects of early working. The instruments for age started to work and years of schooling consist of parental occupations when the individual was 15 years old (z_1), the number of teachers per school for the state where the individual lived when she was 7 and 11 years old and the GDP per capita for the state where the individual lived when she was 12 years old (z_2)[15]. These instruments should be correlated to a person's child labor and schooling decisions, but uncorrelated to her unobservable health endowments, once her demographic characteristics, family background and the current labor market economic conditions are controlled for. Again, I will test the relevance and validity of instruments via tests of excluded instruments in the first stage and overidentification tests in the second stage respectively.

3.4.2 Child Labor and Adult Schooling

In this study, an adult's education level is captured by the years of schooling he obtained. We already notice from Figure 3.4b that the effects of the child labor decision on achieved education level are quite different for different quantiles of years of schooling. Therefore, I estimate quantile regression. The standard linear conditional quantile regression model treats the conditional distribution of the response variable as a linear function of covariates. To be more specific, let $Q_q(sch|x)$ denote the qth standard linear conditional quantile function of the response variable years of schooling given covariates x (including age started to

[15] The procedure to construct instruments and the data source of instruments here are the same as described in the previous section.

work, x_1, z_2 and z_3). Then for the qth quantile (0<q<1), the model can be written as:

$$Q_q(sch_i|x_i) = \pi_{3q} + \alpha_{3q} * startage_i + x_{1i}^{'}\delta_{3q} + z_{2i}^{'}\sigma_{3q} + z_{3i}^{'}\eta_{3q} = x_i^{'}\theta_q \tag{3.3}$$

Note that the parameters θ_q (including π_{3q}, α_{3q}, δ_{3q}, σ_{3q}, η_{3q}) are allowed to vary across quantiles. The qth quantile regression estimator $\hat{\theta}_q$ minimizes over θ_q the objective function

$$\sum_{i:\ y_i \geq x_i^{'}\theta}^{N} q|sch_i - x_i^{'}\theta_q| + \sum_{i:\ y_i < x_i^{'}\theta}^{N} (1-q)|sch_i - x_i^{'}\theta_q|$$

where 0<q<1. In this study, I estimate the schooling equation (3.3) at quantiles 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8 and 0.9.

Although one's early working decision may be correlated to the unexplained components of schooling, I do not have valid instruments for age started to work, hence I will just report results from OLS and quantile regressions.[16]

3.5 Empirical Results

3.5.1 Child Labor and Adult Income

Classical analyses of income models estimate separate models for men and women, since usually there exist substantial gender differences on the wage effects. However, in this study, I test and can not reject the null hypothesis of the equality of all coefficients (except the intercepts) in the male and female models under the 5% significance level, in either the urban or the rural sample. Hence, I pool the men and women samples and include a gender dummy in the income model.

In order to estimate the effects of early labor market entry on current adult earnings, I begin by treating the child labor decision, education and health as exogenous.[17] Table 3.2 reports the OLS coefficient estimates of the income model. The control variables are

[16] I tried the parental occupations when the individual was 15 years old as instruments for the child labor decision in the schooling model, but they cannot pass the overidentification test, indicating the invalidity of these instruments.

[17] I firstly estimate the income equation by quantile regression method with and without considering the endogeneity problem for quantiles 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8 and 0.9. I cannot reject the null hypothesis that the coefficients of age started to work across quantiles are equal to each other. Therefore, OLS and 2SLS coefficient estimates are reported as the final estimates of the income model (3.1).

age started to work, years of schooling, self-reported health status, demographic factors including age, age-squared, gender and race, parental education levels and the GDP per capita of individual's current residence state at the interviewing year. When I treat the individual's child labor decision, education and health conditions as exogenous, whether an individual worked during childhood does not affect her current income, holding other factors constant. This is true for both the urban and rural adults. The more educated the person, the more income she earns. An urban resident's health condition is positively correlated to her earnings.

As described above, the 2SLS method is employed to deal with the possible issues of endogeneity and measurement error. I use parental occupations when the individual was 15 years old, the number of teachers per school of the state where the individual lived when he was 7 and 11 years old and the GDP per capita of the state where the individual lived when he was 12 years old, as well as the number of hospitals, beds and doctors per 1000 inhabitants of the state where the individual lived when he was 7 years old as instruments. The first-stage regression results for the urban and rural samples are presented in Table 3.3a and 3.3b. For the age started to work, years of schooling and dummies for health condition, the F test of excluded instruments all indicate the joint significance of instruments. However, the relative low F-statistics imply that the instruments may not have strong prediction power in explaining endogenous variables. Therefore, the 2SLS estimates may be biased in the direction of OLS estimates due to the possible issue of weak instruments (Bound, Jaeger and Baker, 1995). This is a limitation of the present study. Males enter the labor market about two years earlier, receive one year less of schooling and have higher probability to report excellent health. Compared to white people, urban black and other-raced individuals start working earlier and receive less education, and report almost the same health status. Parental occupations do affect child labor decision: compared to individuals whose parents were employees, individuals with fathers who did not work or were self-employed or mothers who worked without a payment when the individual was 15 years old, enter the labor market at younger ages. This is consistent with our expectation. When father does not work or mother is unsalaried, the household may face a credit constraint and need the child to work to supplement the household income. Besides, when father is self-employed, such as working on the own farm or factory, the child may need to enter the labor market early to help his father.

Table 3.4 presents the second-stage regression results of the income equation. I cannot reject the null hypothesis of overidentification test of all instruments, indicating the validity of my instruments for child labor decision, years of schooling and health condition, for both the urban and rural samples. Early entry into the labor market has no significant impact on adult earnings for urban residents, but has a negative and substantial income impact for rural residents, after controlling for the schooling level and health condition. Entering the labor market one year later increases monthly earnings by 16.7% for a rural resident, which is indeed a sizable effect. An early rural labor market entrant suffers a lower income during adulthood since early working may adversely affect the schooling quality which will in turn impose a negative impact on adult income.

The comparison between the OLS and 2SLS estimates is also of interest. While in neither the OLS nor 2SLS models is the estimated income impact of child labor significantly different from zero for the urban sample, the OLS estimate of starting age lies below the 2SLS estimate for the rural sample. This implies that the possible endogeneity and measurement error bias the effect of early working on adult earnings downward.

The coefficients other than age started to work have the expected signs. For an urban individual, the higher of education level, the more income he earns, and excellent health brings in higher income than poor health. Income rises as he ages, probably owing to the accumulation of working experience, but the return to aging falls. There is a gender gap in earnings: males typically get higher earnings than female workers. When the GDP per capita of the residence state increases indicating a better macroeconomic environment, individuals get higher earnings.

3.5.2 Child Labor and Adult Health

I use model (3.2a) and (3.2b) to estimate the long-term impacts of early entry into the labor force on adult health. The self-assessed health index and the incidence of health problems are the dependent variables, and the controls include the individual's age, age-squared, gender, race, parental education levels, the GDP per capita of the residence state at the interviewing year, the number of health facilities per 1000 inhabitants of the state where the individual lived when she was 7 years old.[18]

I start by estimating the ordered probit model (3.2a) without considering the possible endogeneity problem. The marginal effects rather than the coefficient estimates of the model (3.2a) are reported in Table 3.5. Column 1, 2, and 3 correspond to the marginal effects on the probability that the health index equals 1, 2 and 3 for the urban residents, while column 4, 5 and 6 correspond to the marginal effects for the rural residents. From Table 3.5, there is no significant effect of child labor on adult health for the urban sample, but in the rural areas, an early labor market entrant is less likely to report very good or excellent health, and more likely to report poor, average or good health.

When I take the endogeneity and measurement error issues into account, an IV ordered probit specification of the health model (3.2a) is estimated, with the parental occupations when the individual was 15 years old (z_1), the availability and quality of local education system and the fluctuations of local labor market represented by z_2 being employed to identify the child labor decision and years of schooling. Table 3.6 presents the first-stage regression results. The first two columns of Table 3.6 correspond to the starting age and schooling equations for the urban sample, while the last two columns are the first-stage estimates for the rural sample. From those results, we can find that the instruments are correlated to the age of entry into the labor market and schooling decision and jointly significant.

The second-stage estimates of the relationship between the health index and child work activity are presented in Table 3.7. Column 1, 2 and 3 correspond to the marginal effects on the probability that the health index equals 1, 2 and 3 for the urban residents, and column 4, 5 and 6 correspond to the marginal effects for the rural residents. The comparison between the ordered probit and IV ordered probit estimates indicate that the endogeneity causes the

[18]I separate the health models into men and women and test the equality of all coefficients (except the intercepts) of the male and female models, for the urban and rural samples respectively. I cannot reject the null hypothesis of the equality of all coefficients (except the intercepts) of male and female models for the urban sample under the 5% significance level, indicating that that it's not necessary to separate the urban health model into men and women. And I reject the null for the rural sample. But this is not a strong indication of different models for men and women, since there are only 464 and 202 observations for the rural male and female samples respectively, the rural male and female models are poorly estimated. Hence, I pool men and women for the rural health model for simplicity.

estimated effect of child labor to be smaller than it really is. Results from Table 3.7 suggest that there is no significant health consequence for an urban early labor market entrant, but there exists a significant negative effect of early working on rural adult health: the probability of reporting very good or excellent health falls and the probability of reporting poor or average health rises as one enters the labor force earlier.

One thing worthy of notice is the opposite effects of schooling on the health status of urban and rural adults. An increase in the schooling years benefits urban adult health but harms rural adult health, due to the twofold impacts of education. On the one hand, as one achieves higher level of education, she would gain access to more knowledge about health care, and usually would take care of herself more carefully, which is expected to have a positive impact on health condition. On the other hand, the more educated of an individual, the more likely she would pay attention to her own health condition and to recognize and report health problems, which suggests the negative effect of schooling on reported health condition. Back to my sample, from the descriptive analysis in Section 3.3, urban residents receive 3.5 more years of schooling than rural residents on average. Hence, when rural residents are relatively low educated on average, one additional year of schooling may be more effective on recognizing and reporting health problems leading to a negative effect of schooling on reported health condition, while the average urban residents finish the upper primary education, the impact of schooling may focus more on getting access to knowledge about health care and consequently has a positive effect on reported health. Additionally, the probability to report very good or excellent health decreases as persons age and urban men report being healthier than women.

An additional probit model with the incidence of health problems as the dependent variable is estimated. The probit estimates, the first-stage and second-stage regression results of the IV probit specification of model (3.2b) are presented in Table 3.8, 3.9 and 3.10 respectively. The main results from health model (3.2b) are consistent with those derived from model (3.2a): child labor is associated with worse adult health. However, although Table 3.7 shows child labor only affects future health in the rural sample, Table 3.10 exhibits significant negative linkages between early working and adult health for both the urban and rural samples. One year earlier of entry into the labor force leads to an increase of 2.1 and

9.1 percentage points in the probability of getting health problems in adulthood, for an urban and rural resident respectively. The probit estimates of child labor are centered over zero for both urban and rural samples, compared to the IV estimates. Again, this implies the existence of endogeneity. Urban males are about 9 percentage points less likely to get health problems than females.

While it is clear that an early labor market entrant suffers worse health outcomes, I find big area differences of child labor effects: whether an individual worked as a child laborer does not affect how he evaluates his health condition in urban areas but does in rural areas; meanwhile, as one enters the labor market one year earlier, the probability of getting health problems increases by 2.1 and 9.1 percentage points for an urban and rural resident respectively.

These substantial area differences in the health impacts of early exposure to work may be due to the different working environments and conditions for the first job in the urban and rural samples. In the sample, among those rural residents who started to work before 18 years old, more than 70% were employed in the agriculture sector, while among those urban residents who entered the labor market as a child, most of them worked in the service (24%), retailing (15%), manufacturing and construction industry (19%). As one may notice that the agriculture sector is ranked as one of the most hazardous sectors in terms of morbidity and mortality (Fassa et al., 2000). Fassa et al. (2000) points out that children employed in the agriculture sector are easily injured by dangerous machinery, exposure to strenuous labor, chemicals and adverse weather (e.g. heat). Also, agriculture is among the less regulated sectors where the laws protecting children are very difficult to enforce. In contrast, child workers in the urban sample who are involved in manufacturing, retailing, services and other industries may not suffer from as adverse working conditions and poorly regulated working environments as do their rural peers.

3.5.3 Child Labor and Adult Schooling

From Figure 3.4b, the distributions of years of schooling conditional on starting ages are quite different. Consequently, for both the urban and rural samples, I use the quantile regression approach to capture the associations between early working and schooling. Model

(3.3) is employed to explore the relationship between child labor and adult schooling.[19] I report OLS estimates first, followed by quantile regression results. A major difference between OLS and quantile regression is that OLS characterizes the mean of the distribution whereas the quantile regression explores the full shape of the conditional distribution of the dependent variable. Relatively speaking, quantile regression provides a more precise estimation and a more complete picture of the conditional distribution of years of schooling.

As explained above, although there may exist some unobservable factors affecting both child labor decision and educational attainment, I do not have suitable instruments that are sufficiently correlated to the child labor decision but not correlated to the unexplained components of schooling, hence my estimates of the impacts of child labor on adult schooling are suggestive but not causal.

Table 3.11a and 3.11b exhibit the OLS estimates first, followed by the quantile regression estimates of the schooling model at the 10th, 20th, 30th, 40th, 50th, 60th, 70th, 80th and 90th conditional percentiles, for the urban and rural samples respectively. Of great interest are the coefficients on the age an individual started to work. These parameters estimate the changes in specific conditional percentiles of years of schooling caused by one unit change in the starting age. The child labor decision imposes significant negative effects on an adult's educational attainment for almost all the quantiles. The later one enters the labor market, the more years of schooling she attains. The marginal changes associated with one year later of entry into the work force in the median conditional quantile of years of schooling are an increase of 0.284 years and 0.178 years, for the urban and rural samples respectively. The coefficients of age started to work vary considerably across quantiles. For instance, there is an 80 percent difference between the starting age coefficients for the 0.5 quantile and 0.1 quantile in the urban sample (the coefficient estimates for starting age for the 0.5 quantile and 0.1 quantile are statistically different from each other (p=0.000)), while in

[19] I separate the schooling quantile regression models into men and women for the urban and rural samples, and test the equality of all coefficients (except for the intercepts) of the male and female models. In most (except for the 30th quantile in the urban sample) quantile regressions for either the urban or the rural sample, I cannot reject the null hypothesis of the equality of all coefficients (except the intercepts) of the male and female models.

I then estimate the urban schooling model for the 30th quantile, with the sample being separated into men and women. The schooling effects of early working from the male and female models are quite close to each other. Hence, I pool men and women for the schooling model for simplicity, for both the urban and rural samples.

the rural sample the starting age coefficient for the 0.7 quantile is close to 134 percent above that of the 0.2 quantile (the coefficient estimates for starting age for the 0.7 quantile and 0.2 quantile are also statistically different from each other (p=0.006)). I also test and reject the equality of coefficients of age started to work across quantiles. Most noticeably, starting age has much greater impacts at the middle conditional quantile of schooling for the urban sample and at 0.7 quantile of schooling for the rural sample than those in tails of the schooling distribution. As a matter of fact, quantile regression estimates exhibit an inverse U-shaped trend which rises over the percentiles until around middle quantile and then falls for the urban sample, suggesting that the schooling effect of early working is greater for an individual with middle level of education than one in the tails of the schooling distribution. A similar trend can be found in the rural sample estimates.

Another thing worthy of notice is that child labor imposes greater effects on schooling for urban residents than rural residents (except for the 90th percentile). For some quantiles of years of schooling (e.g., the 10th and 20th quantiles), the impact of child labor on urban residents is twice or even more than twice that on rural residents. Similar to the health model, this large area difference may be due to the different working environment and conditions for the first job in the urban and rural samples. Most rural residents were involved in agricultural work in the first job, and agricultural work is often seasonal work and may be more compatible with schooling than working in urban areas.

Clearly, the quantile regression estimates are different from the OLS estimates. According to the linear regression model, an urban individual's schooling level would increase by 0.227 years if he started working one year later. However, the quantile regression results indicate larger impacts of child labor on the 30th, 40th, 50th, 60th and 70th quantiles of schooling years for the urban sample. For instance, entering the labor market one year later causes the 50th conditional quantile of schooling to increase about 0.284 years for a urban resident. Similar results can be found in the rural sample: the linear regression model underestimates the effects of child labor at the 40th, 50th, 60th, 70th, 80th and 90th quantiles of years of schooling.

Male workers typically receive less schooling than female workers. Urban black and other-raced residents' educational attainments are lower than those of white people. Parental education levels impose positive effects on the individual's schooling level, the more educated of parents, the higher schooling level the individual gets.

3.5.4 Robustness Analysis

All the above results indicate the adverse long-term consequences of early working on adult earnings, health and educational attainment. However, all the analyses are based on the working sample with valid income information, and therefore may be potentially biased. In this section, I re-estimate the health model and the schooling model using the full sample including all individuals with valid and non-valid (missing or zero) income data to work as a robustness check.[20]

Look at the health model (3.2a) first. The ordered probit estimates and the first-stage regression estimates for the IV ordered probit specification with the full sample can be found in Appendix Table A.3.1 and A.3.2. My instruments are jointly significant in the first-stage regressions. I report the second-stage regression results for the IV ordered probit specification in Table 3.12. Age started to work cannot be statistically differentiated from zero in the urban full sample. And in the rural case, one year later of entering the labor market reduces the probability for an individual to assess her health as "good" at a very small magnitude but increases her probability to report "very good or excellent" health significantly. This is consistent with our main finding from Table 3.7: early working does not affect an urban resident's health but imposes an adverse health effect on a rural resident.

Table 3.13 presents the second-stage regression results from the full sample for the IV probit specification of the health model (3.2b).[21] My instruments are jointly significant in the first stage and pass overidentification tests in the second stage. Again, early working exhibits great adverse health effects: as one enters the labor market one year earlier, the probability of reporting health problems in adulthood rises by 1.9 and 5.7 percentage points for an urban and rural individual respectively. A great difference of child labor effects between the urban and the rural sample shows up again: while early working does not

[20] My income model is probably subjected to the sample selection bias. However, sample selection is not the main research problem of interest in the present study, so I do not correct for it here.

[21] The probit estimates and the first-stage regression results for the IV probit specification of the health model are shown in Appendix Table A.3.3 and Table A.3.4.

affect how one evaluates his health condition in urban areas but does in rural areas, one year earlier entry into the labor market increases the probability of reporting health problems much more greatly for a rural resident than an urban resident.

Table 3.14a and 3.14b exhibit the OLS and quantile regression results from the full sample for the schooling model. A comparison between Table 3.11 and 3.14 shows that the coefficients of age started to work estimated from the full urban sample are close to those estimated from the working urban sample, and the coefficients of age started to work have the same signs for the working and full sample in the rural case, though there exist some differences in magnitudes. The fact that the adverse schooling impacts of child labor are greater on urban adults than their rural peers is true for most quantiles of schooling (except for the 90th percentile) in both the working and the full samples.

In either the health model or the schooling model, most coefficients other than age started to work have the same signs for the working sample and the full sample, and for those estimates which have opposite signs, most of them are insignificant, although there exist some differences in the magnitudes of the coefficients of the working and full samples, especially in the rural case.

In sum, the story we get from the full sample is consistent with the one from the working sample: child labor negatively affect adult health and schooling, and early entry into the labor market imposes different effects on the urban and rural residents.

3.6 Discussion and Policy Implications

This study investigates in great detail the long-run effects of working as a child laborer on an individual's health and socioeconomic conditions. It explores the Brazilian PPV data set and analyzes the long-term income, health and schooling effects of early working for the urban and rural samples separately. In order to deal with the possible endogeneity and measurement error problems, I employ instruments to estimate the income model and the health model (3.2a) and (3.2b), with the 2SLS method used for the income model, the IV ordered probit method used for the health model (3.2a) and the IV probit method used for the health model (3.2b). However, due to the fact that it is too difficult to find a suitable instrument which can decompose the effect of child labor decision from schooling decision, my work does not take into account of the endogeneity problem in the schooling model. The quantile regression technology is used to capture the different effects of early working on schooling across quantiles. It would be informative to find a valid instrument for child labor decision in the schooling model, to examine the causal relationship between early working and adult schooling.

The results presented in this study suggest that early exposure to work for a rural resident leads to lower earnings when controlling for schooling and health conditions, and a worse self-assessed health index when controlling for schooling. Also, an urban/rural adult has a higher probability to get health problems if she worked during childhood. As for schooling, the later one enters the labor market, the more years of schooling she obtains. While both the income and health effects of child labor on rural residents are greater than those on urban residents, urban residents suffer greater adverse schooling impacts than their rural peers. Although early working may help young laborers to accumulate working experience and finance the household or schooling, the combined final effects of child labor on a person's future development are negative, i.e., a child who starts to work early suffers adverse health and socioeconomic consequences in the long run. The findings for the health and schooling models are robust when the full sample including all individuals with valid and non-valid (missing or zero) income information are used for estimation.

My findings have important implications: all the aforementioned negative effects of child labor on adult outcomes make a strong call to reduce child labor in Brazil and other developing countries. In addition, the different effects of early working on urban and rural adults should be taken into account when child labor policies are proposed. Given that rural children are more vulnerable to the adverse consequences of early working at many aspects, we should pay special attention to tackle the issue of child labor in the rural area.

Table 1.1: Summary Statistics

	Sample 1 (N=2,105)		Sample 2 (N=1,928)		Sample 3 (N=1,618)	
Variables	Mean	Std. Dev	Mean	Std. Dev	Mean	Std. Dev
Dependent Variables						
employed	0.9226	0.2673				
occupation			2.5737	1.2334		
ln(earnings)					6.4704	0.7433
The Big Five Personality Traits						
extraversion	31.6822	7.4875	31.7863	7.5017	31.5797	7.5619
agreeableness	38.4033	5.6526	38.3729	5.6442	38.4302	5.5484
conscientiousness	36.9748	5.9133	37.1385	5.8255	37.0884	5.8462
emotional stability	33.1468	8.0110	33.5124	7.7905	33.4629	7.7554
imagination	35.5929	5.8160	35.5934	5.7953	35.5779	5.8713
Other Explanatory Variables						
married	0.7121	0.4529	0.7334	0.4423	0.7398	0.4389
urban	0.7169	0.4506	0.7116	0.4531	0.7126	0.4527
wales	0.0542	0.2264	0.0534	0.2249	0.0544	0.2269
scotland	0.1192	0.3241	0.1162	0.3205	0.1199	0.3249
england	0.8266	0.3787	0.8304	0.3754	0.8257	0.3795
no academic qualification	0.1629	0.3694	0.1457	0.3529	0.1428	0.3499
compulsory education	0.4869	0.4999	0.4870	0.5000	0.4790	0.4997
extended secondary education	0.0931	0.2907	0.0991	0.2988	0.1001	0.3003
tertiary education	0.2570	0.4371	0.2682	0.4431	0.2781	0.4482
reading score at age 11	17.4485	6.0289	17.6753	5.9578	17.7027	5.9587
father's years of schooling	10.0846	2.1638	10.1328	2.1994	10.1502	2.2534
Instruments						
At Age 11						
hostility	0.7739	1.7011	0.7137	1.5881	0.6557	1.4673
immaturity	0.5059	0.9214	0.4829	0.8990	0.4883	0.9057
nervous symptoms	0.1159	0.3812	0.1115	0.3736	0.0995	0.3400
At Age 7						
having difficulty concentrating	0.3069	0.4613	0.3050	0.4605	0.3066	0.4612
preferring to do things alone	0.6703	0.4702	0.6748	0.4686	0.6743	0.4688
being bullied	0.3530	0.4780	0.3454	0.4756	0.3517	0.4776
being fidgety	0.4399	0.4965	0.4419	0.4967	0.4425	0.4968
worrying about many things	0.4651	0.4989	0.4663	0.4990	0.4660	0.4990
being irritable	0.4793	0.4997	0.4735	0.4994	0.4802	0.4998
being upset by new situation	0.2936	0.4555	0.2982	0.4576	0.3059	0.4609
biting nails	0.2057	0.4043	0.2054	0.4041	0.2015	0.4012
being disobedient	0.6043	0.4891	0.6017	0.4897	0.6020	0.4896

Note: Sample 1 is for the employment status model, Sample 2 is for the occupational choice model, Sample 3 is for the earnings model.

Table 1.2: Probit and IVprobit Estimates for Employment Status Model

Variables	Probit (b)	Probit (b')	IVprobit (b)	IVprobit (b')
The Big Five				
extraversion	0.0009	0.0066	-0.0057	-0.0425
	(0.0009)		(0.0052)	
agreeableness	-0.0023**	-0.0128	0.0114**	0.0646
	(0.0011)		(0.0056)	
conscientiousness	0.0029***	0.0171	0.0137***	0.0809
	(0.0010)		(0.0034)	
emotional stability	0.0029***	0.0235	0.0013	0.0103
	(0.0007)		(0.0043)	
imagination	-0.0044***	-0.0257	-0.0442***	-0.2572
	(0.0012)		(0.0058)	
Other Explanatory Variables				
married	0.0652***	0.0295	0.0380*	0.0172
	(0.0109)		(0.0228)	
urban	-0.0176	-0.0079	-0.0038	-0.0017
	(0.0129)		(0.0169)	
wales	-0.0110	-0.0025	0.0287	0.0065
	(0.0239)		(0.0333)	
scotland	-0.0221	-0.0072	-0.0503**	-0.0163
	(0.0157)		(0.0197)	
compulsory education	0.0347**	0.0173	0.0343	0.0171
	(0.0135)		(0.0210)	
extended secondary education	0.1002***	0.0291	0.1658***	0.0482
	(0.0282)		(0.0336)	
tertiary education	0.0715***	0.0313	0.1700***	0.0743
	(0.0196)		(0.0282)	
reading score at age 11	0.0035***	0.0213	0.0139***	0.0837
	(0.0011)		(0.0020)	
father's years of schooling	0.0085**	0.0183	0.0095**	0.0206
	(0.0043)		(0.0047)	
Observations	2,105		2,105	
Overidentification test of all instruments				
			$\chi^2(7) = 4.365$	
P-value			0.7369	

Notes: Marginal effects are reported rather than probit coefficients. Heteroskedasticity-robust standard errors in parentheses. *** Significant at the 1% level, ** Significant at the 5% level, * Significant at the 10% level. The Amemiya-Lee-Newey statistics for overidentification test of instruments are reported. The coefficient b' represents the change in employment probability for a one standard deviation increase in the independent variable ($b'_x = b_x * \sigma_x$).

Table 1.3: IVprobit Estimates - First-stage Regressions of Employments Status Model

Variables	extraversion	agreeableness	conscientiousness	emotional stability	imagination
married	1.2002***	0.1338	0.8251***	1.2775***	-0.4965*
	(0.3587)	(0.2679)	(0.2962)	(0.3955)	(0.2537)
urban	0.7650**	0.4833*	0.0095	-0.6427*	0.3657
	(0.3589)	(0.2762)	(0.2876)	(0.3804)	(0.2632)
wales	1.1057	0.0215	-0.3459	1.4732**	0.6852
	(0.6864)	(0.5892)	(0.6120)	(0.6980)	(0.4996)
scotland	-0.4840	0.8841**	-0.4699	0.5708	-0.5322
	(0.4521)	(0.3721)	(0.3855)	(0.5356)	(0.3257)
compulsory education	1.6102***	0.9432***	0.9983**	1.5100***	0.4413
	(0.4601)	(0.3659)	(0.3947)	(0.5266)	(0.3383)
extended secondary education	1.7834***	0.5829	1.5031***	2.0102***	2.1800***
	(0.6790)	(0.5377)	(0.5502)	(0.7649)	(0.4839)
tertiary education	2.7363***	1.7994***	1.3578***	2.2567***	3.1569***
	(0.5869)	(0.4500)	(0.4851)	(0.6452)	(0.4316)
reading score at age 11	-0.0258	0.0474**	-0.0279	0.0041	0.2804***
	(0.0313)	(0.0239)	(0.0261)	(0.0344)	(0.0226)
father's years of schooling	0.0906	0.0719	0.1130**	-0.0029	0.0704
	(0.0793)	(0.0567)	(0.0574)	(0.0851)	(0.0497)
Instruments					
At Age 11					
hostility	0.1338	-0.1063	-0.0833	-0.3388***	0.0581
	(0.0892)	(0.0798)	(0.0777)	(0.1088)	(0.0619)
immaturity	-0.5172***	-0.3440**	-0.4596***	-0.0057	-0.1003
	(0.1647)	(0.1413)	(0.1496)	(0.1923)	(0.1115)
nervous symptoms	0.3645	-0.3623	0.0454	0.1274	-0.1382
	(0.3877)	(0.3209)	(0.3521)	(0.4643)	(0.2683)
At Age 7					
having difficulty concentrating	0.5764*	0.2935	-0.0309	0.0872	-0.0506
	(0.3495)	(0.2892)	(0.2906)	(0.3964)	(0.2143)
preferring to do things alone	-0.5295	0.0330	0.3848	-0.7463**	0.1155
	(0.3335)	(0.2575)	(0.2745)	(0.3720)	(0.1985)
being bullied	-0.7043**	0.2032	-0.1089	-0.4295	0.5416***
	(0.3492)	(0.2635)	(0.2806)	(0.3791)	(0.2068)
being fidgety	-0.0385	-0.0991	-0.4938*	-0.2672	-0.4994***
	(0.3309)	(0.2637)	(0.2762)	(0.3705)	(0.1904)
worrying about many things	0.6647**	0.3983	0.5775**	-0.3664	0.2946
	(0.3289)	(0.2622)	(0.2734)	(0.3736)	(0.1951)
being irritable	0.4591	0.0135	0.5220*	-0.3533	0.2741
	(0.3194)	(0.2588)	(0.2667)	(0.3618)	(0.1924)
being upset by new situation	-0.8396**	-0.3690	0.1327	-0.8403**	-0.6417***
	(0.3734)	(0.2872)	(0.2976)	(0.4122)	(0.2377)
biting nails	-0.2235	-0.4831	0.1265	-0.0347	-0.2896
	(0.3846)	(0.3126)	(0.3172)	(0.4364)	(0.2370)
being disobedient	-0.1823	-0.4576*	-0.4006	-0.0395	-0.1451
	(0.3287)	(0.2680)	(0.2772)	(0.3780)	(0.1988)
Constant	28.5936***	35.7771***	34.7855***	32.6158***	28.9348***
	(1.0856)	(0.8119)	(0.8594)	(1.1803)	(0.7004)
Observations	2,105	2,105	2,105	2,105	2,105
Joint test of instruments					
$\chi^2(12)$	32.16	24.41	31.80	30.57	26.44
P-value	0.0013	0.0179	0.0015	0.0023	0.0093

Note: Heteroskedasticity-robust standard errors in parentheses. *** Significant at the 1% level, ** Significant at the 5% level, * Significant at 10% level.

Table 1.4: MNL Results for Occupational Choice Model

Variables	Managerial (b)	b'	Non-manual (b)	b'	Professional (b)	b'	Manual (b)	b'
The Big Five								
extraversion	0.0062***	0.0467	-0.0006	-0.0041	-0.0012	-0.0090	-0.0045***	-0.0335
	(0.0015)		(0.0015)		(0.0009)		(0.0015)	
agreeableness	-0.0064***	-0.0364	0.0055***	0.0313	0.0008	0.0044	0.0001	0.0007
	(0.0019)		(0.0021)		(0.0013)		(0.0020)	
conscientiousness	0.0042**	0.0247	-0.0029	-0.0167	0.0033***	0.0190	-0.0046**	-0.0270
	(0.0018)		(0.0019)		(0.0011)		(0.0018)	
emotional stability	-0.0002	-0.0019	-0.0011	-0.0084	-0.0004	-0.0030	0.0017	0.0132
	(0.0013)		(0.0014)		(0.0008)		(0.0014)	
imagination	0.0020	0.0113	0.0012	0.0068	-0.0006	-0.0036	-0.0025	-0.0146
	(0.0021)		(0.0021)		(0.0013)		(0.0021)	
Other Explanatory Variables								
married	0.1273***	0.0563	-0.0427*	-0.0189	0.0152	0.0067	-0.0998***	-0.0442
	(0.0239)		(0.0222)		(0.0145)		(0.0216)	
urban	-0.0656***	-0.0297	0.0496**	0.0225	-0.0216*	-0.0098	0.0376*	0.0171
	(0.0208)		(0.0222)		(0.0126)		(0.0222)	
wales	-0.1314**	-0.0296	0.0734*	0.0165	-0.0267	-0.0060	0.0847*	0.0191
	(0.0517)		(0.0424)		(0.0303)		(0.0452)	
scotland	-0.0107	-0.0034	-0.0244	-0.0078	-0.0220	-0.0071	0.0571*	0.0183
	(0.0313)		(0.0326)		(0.0217)		(0.0294)	
compulsory education	0.0358	0.0179	0.0379	0.0190	-0.0193	-0.0097	-0.0544**	-0.0272
	(0.0346)		(0.0360)		(0.0264)		(0.0270)	
extended secondary education	0.1322***	0.0395	0.1438***	0.0430	-0.0236	-0.0071	-0.2524***	-0.0754
	(0.0429)		(0.0439)		(0.0333)		(0.0421)	
tertiary education	0.0775**	0.0344	0.1560***	0.0691	0.0780***	0.0346	-0.3116***	-0.1381
	(0.0391)		(0.0385)		(0.0256)		(0.0352)	
reading score at age 11	0.0029	0.0174	0.0083***	0.0493	0.0035***	0.0209	-0.0147***	-0.0877
	(0.0020)		(0.0020)		(0.0012)		(0.0019)	
father's years of schooling	0.0097**	0.0214	-0.0025	-0.0056	0.0081***	0.0177	-0.0152**	-0.0335
	(0.0047)		(0.0049)		(0.0020)		(0.0061)	
Observations	1,928		1,928		1,928		1,928	

Note: Marginal effects are reported rather than multinomial logit coefficients here. Heteroskedasticity-robust standard errors in parentheses. *** Significant at the 1% level, ** Significant at the 5% level, * Significant at the 10% level. The coefficient b' represents the change in the probability of being in each occupation for a one standard deviation increase in the independent variable ($b'_x = b_x * \sigma_x$). The detailed information about each occupation group can be found in Appendix Table A.1.1.

Table 1.5: Two-stage MNL Estimates - First-stage Regressions of Occupational Choice Model

Variables	extraversion	agreeableness	conscientiousness	emotional stability	imagination
married	1.1741***	0.2150	0.6038*	0.8351**	-0.2615
	(0.3872)	(0.2825)	(0.3094)	(0.4080)	(0.2671)
urban	0.7810**	0.5357*	0.1525	-0.4687	0.4693*
	(0.3746)	(0.2846)	(0.2979)	(0.3922)	(0.2727)
wales	1.1071	0.1231	-0.3641	1.4499**	0.8928*
	(0.7509)	(0.6405)	(0.6587)	(0.7365)	(0.5343)
scotland	-0.7147	1.0190***	-0.6128	0.2110	-0.7365**
	(0.4814)	(0.3780)	(0.3949)	(0.5666)	(0.3331)
compulsory education	1.5006***	1.0474***	1.0592**	1.2471**	0.6532*
	(0.4936)	(0.3972)	(0.4296)	(0.5487)	(0.3614)
extended secondary education	1.8709***	0.7550	1.5688***	1.8388**	2.4482***
	(0.7028)	(0.5674)	(0.5823)	(0.7803)	(0.5026)
tertiary education	2.6939***	1.9561***	1.4989***	1.9283***	3.4366***
	(0.6196)	(0.4767)	(0.5083)	(0.6600)	(0.4500)
reading score at age 11	-0.0427	0.0425*	-0.0357	-0.0040	0.2694***
	(0.0324)	(0.0251)	(0.0273)	(0.0348)	(0.0238)
father's years of schooling	0.1085	0.0974*	0.0708	-0.0335	0.0748
	(0.0815)	(0.0575)	(0.0582)	(0.0862)	(0.0505)
Instruments					
At Age 11					
hostility	0.1546	-0.1156	-0.0956	-0.2916**	-0.0106
	(0.0988)	(0.0885)	(0.0862)	(0.1179)	(0.0747)
immaturity	-0.4623**	-0.2714*	-0.4817***	0.0204	-0.1315
	(0.1878)	(0.1468)	(0.1611)	(0.2001)	(0.1526)
nervous symptoms	0.4396	-0.5516*	-0.1453	-0.1154	-0.1679
	(0.4397)	(0.3331)	(0.3701)	(0.4816)	(0.3131)
At Age 7					
having difficulty concentrating	0.4021	0.1940	-0.0627	-0.0806	-0.2388
	(0.3826)	(0.2958)	(0.3015)	(0.3982)	(0.2692)
preferring to do things alone	-0.4565	-0.0628	0.2705	-0.8339**	0.2449
	(0.3694)	(0.2670)	(0.2834)	(0.3756)	(0.2565)
being bullied	-0.7122*	0.1474	0.0047	-0.0485	0.4627*
	(0.3793)	(0.2783)	(0.2927)	(0.3872)	(0.2563)
being fidgety	0.0918	-0.2003	-0.5764**	-0.3167	-0.3154
	(0.3556)	(0.2713)	(0.2861)	(0.3770)	(0.2511)
worrying about many things	0.5764	0.4859*	0.4708*	-0.7054*	0.2353
	(0.3604)	(0.2734)	(0.2835)	(0.3808)	(0.2509)
being irritable	0.5011	-0.0236	0.2922	-0.5124	0.3693
	(0.3493)	(0.2706)	(0.2796)	(0.3737)	(0.2459)
being upset by new situation	-0.9510**	-0.3520	0.0319	-0.8980**	-0.8365***
	(0.3929)	(0.2887)	(0.3048)	(0.4157)	(0.2683)
biting nails	-0.1317	-0.3630	0.1384	0.2079	-0.4193
	(0.4255)	(0.3260)	(0.3324)	(0.4423)	(0.2967)
being disobedient	-0.2828	-0.4894*	-0.3184	0.0768	-0.0878
	(0.3626)	(0.2793)	(0.2869)	(0.3860)	(0.2583)
Constant	28.8395***	35.4153***	35.7215***	33.9412***	28.5682***
	(1.1497)	(0.8559)	(0.8984)	(1.2194)	(0.7643)
Observations	1,928	1,928	1,928	1,928	1,928
R-squared	0.0330	0.0425	0.0270	0.0295	0.2149
Joint test of instruments					
$F(12, 1906)$	2.04	1.86	2.09	2.60	1.91
P-value	0.0183	0.0349	0.0149	0.0020	0.0290

Notes: Heteroskedasticity-robust standard errors in parentheses. *** Significant at 1% level, ** Significant at 5% level, * Significant at 10% level.

Table 1.6: Two-stage MNL Results for Occupational Choice Model

Variables	Managerial (b)	b'	Non-manual (b)	b'	Professional (b)	b'	Manual (b)	b')
The Big Five								
extraversion	0.0402**	0.3018	-0.0283*	-0.2126	-0.0078	-0.0588	-0.0041	-0.0304
	(0.0169)		(0.0146)		(0.0097)		(0.0130)	
agreeableness	0.0415	0.2343	0.0172	0.0973	-0.0335*	-0.1893	-0.0252	-0.1423
	(0.0339)		(0.0286)		(0.0177)		(0.0230)	
conscientiousness	0.0184	0.1071	-0.0338	-0.1970	0.0153	0.0892	0.0001	0.0007
	(0.0280)		(0.0242)		(0.0147)		(0.0199)	
emotional stability	-0.0119	-0.0924	-0.0104	-0.0808	0.0048	0.0372	0.0175	0.1360
	(0.0148)		(0.0128)		(0.0080)		(0.0108)	
imagination	-0.0500*	-0.2898	0.0493**	0.2857	0.0082	0.0475	-0.0075	-.0434
	(0.0278)		(0.0238)		(0.0155)		(0.0203)	
Other Explanatory Variables								
married	0.0587	0.0260	0.0317	0.0140	0.0227	0.0100	-0.1131***	-0.0500
	(0.0470)		(0.0384)		(0.0254)		(0.0369)	
urban	-0.1019***	-0.0462	0.0426	0.0193	-0.0005	-0.0002	0.0598**	0.0271
	(0.0374)		(0.0362)		(0.0196)		(0.0302)	
wales	-0.1073	-0.0241	0.0663	0.0149	-0.0280	-0.0063	0.0689	0.0155
	(0.0812)		(0.0699)		(0.0943)		(0.0652)	
scotland	-0.0645	-0.0207	-0.0347	-0.0111	0.0218	0.0070	0.0774	0.0248
	(0.0674)		(0.0599)		(0.0400)		(0.0485)	
compulsory education	-0.0347	-0.0174	0.0828	0.0414	-0.0002	-0.0001	-0.0479	-0.0239
	(0.0599)		(0.0514)		(0.0396)		(0.0438)	
extended secondary education	0.1544	0.0461	0.1386*	0.0414	-0.0379	-0.0113	-0.2551***	-0.0762
	(0.0943)		(0.0786)		(0.0575)		(0.0710)	
tertiary education	0.0728	0.0323	0.1059	0.0469	0.1044*	0.0463	-0.2831***	-0.1255
	(0.1078)		(0.0894)		(0.0599)		(0.0797)	
reading score at age 11	0.0164**	0.0975	-0.0070	-0.0416	0.0027	0.0162	-0.0121**	-0.0721
	(0.0078)		(0.0070)		(0.0045)		(0.0059)	
father's years of schooling	0.0039	0.0086	-0.0020	-0.0044	0.0113***	0.0248	-0.0132*	-0.0290
	(0.0073)		(0.0069)		(0.0036)		(0.0072)	
Observations	1,928		1,928		1,928		1,928	

Notes: Marginal effects are reported rather than multinomial logit coefficients. Bootstrapped standard errors in parentheses. *** Significant at the 1% level, ** Significant at the 5% level, * Significant at the 10% level. The coefficient b' represents the change in the probability of being in each occupation for a one standard deviation increase in the independent variable ($b'_x = b_x * \sigma_x$). The detailed information about each occupation group can be found in Appendix Table A.1.1.

Table 1.7: OLS and 2SLS Regression Results for Earnings Model

Variables	OLS (b)	OLS (b')	2SLS (b)	2SLS (b')
The Big Five				
extraversion	0.0107***	0.0807	0.0354	0.2675
	(0.0027)		(0.0252)	
agreeableness	-0.0106***	-0.0590	0.0393	0.2178
	(0.0032)		(0.0585)	
conscientiousness	0.0174***	0.1017	0.0579	0.3385
	(0.0036)		(0.0444)	
emotional stability	0.0023	0.0178	0.0211	0.1637
	(0.0024)		(0.0243)	
imagination	0.0005	0.0030	-0.0429	-0.2518
	(0.0037)		(0.0386)	
Other Explanatory Variables				
married	0.2256***	0.0990	0.1322**	0.0580
	(0.0381)		(0.0626)	
urban	-0.0590	-0.0267	-0.0819	-0.0371
	(0.0409)		(0.0587)	
wales	-0.2524***	-0.0573	-0.3099***	-0.0703
	(0.0629)		(0.1173)	
scotland	-0.0375	-0.0122	-0.0838	-0.0272
	(0.0525)		(0.1141)	
compulsory education	0.0435	0.0217	-0.1224	-0.0612
	(0.0480)		(0.1053)	
extended secondary education	0.1998**	0.0600	0.0562	0.0169
	(0.0886)		(0.1789)	
tertiary education	0.3501***	0.1569	0.2128	0.0954
	(0.0618)		(0.1642)	
reading score at age 11	0.0141***	0.0840	0.0265**	0.1580
	(0.0034)		(0.0120)	
father's years of schooling	0.0308***	0.0693	0.0224**	0.0505
	(0.0078)		(0.0099)	
Constant	4.9952***		1.8010	
	(0.2545)		(1.5509)	
Observations	1,618		1,618	
R-squared	0.1648		-	
Overidentification test of all instruments				
Hansen J-statistic			3.873	
P-value			0.7943	

Note: Heteroskedasticity-robust standard errors in parentheses. *** Significant at the 1% level, ** Significant at the 5% level, * Significant at 10% level. R^2 is not reported for the 2SLS specification since it is negative. The coefficient b' represents the percentage change in earnings for a one standard deviation increase in the independent variable ($b'_x = b_x * \sigma_x$).

Table 1.8: 2SLS Estimates - First-stage Regressions of Earnings Model

Variables	extraversion	agreeableness	conscientiousness	emotional stability	imagination
married	1.3170***	0.2945	0.3116	0.6306	-0.1607
	(0.4282)	(0.3073)	(0.3423)	(0.4439)	(0.2938)
urban	0.9048**	0.3882	0.2474	-0.4264	0.4546
	(0.4191)	(0.3086)	(0.3285)	(0.4247)	(0.2999)
wales	1.2324	0.4609	0.1315	2.4351***	1.1110*
	(0.8132)	(0.6554)	(0.7073)	(0.7830)	(0.5828)
scotland	-0.9031*	0.9400**	-0.6254	0.4228	-0.7939**
	(0.5202)	(0.3918)	(0.4219)	(0.6010)	(0.3569)
compulsory education	1.4528***	0.9695**	1.5396***	1.6707***	0.5440
	(0.5435)	(0.4246)	(0.4792)	(0.5975)	(0.4033)
extended secondary education	2.5164***	1.0001	1.7913***	2.8906***	2.4875***
	(0.7867)	(0.6122)	(0.6360)	(0.8448)	(0.5599)
tertiary education	2.7679***	1.8606***	1.8150***	2.4114***	3.4817***
	(0.6810)	(0.5102)	(0.5618)	(0.7078)	(0.4963)
reading score at age 11	-0.0357	0.0349	-0.0459	-0.0206	0.2826***
	(0.0364)	(0.0272)	(0.0304)	(0.0383)	(0.0265)
father's years of schooling	0.0951	0.0945	0.0989	-0.0346	0.0683
	(0.0895)	(0.0615)	(0.0642)	(0.0929)	(0.0556)
Instruments					
At Age 11					
hostility	0.2162*	-0.0820	-0.1320	-0.1931	-0.0366
	(0.1249)	(0.0976)	(0.1008)	(0.1453)	(0.0932)
immaturity	-0.5456***	-0.3566**	-0.5084***	0.0410	-0.1923
	(0.2054)	(0.1574)	(0.1751)	(0.2109)	(0.1696)
nervous symptoms	0.8088	-0.1880	0.0256	-0.4339	0.0944
	(0.5537)	(0.4261)	(0.4727)	(0.5395)	(0.4087)
At Age 7					
having difficulty concentrating	0.4417	0.1523	0.0183	-0.1671	-0.2814
	(0.4208)	(0.3162)	(0.3288)	(0.4309)	(0.2937)
preferring to do things alone	-0.3791	-0.0247	0.3766	-0.5838	0.2911
	(0.4067)	(0.2905)	(0.3120)	(0.4058)	(0.2804)
being bullied	-1.0490**	0.1262	0.2067	-0.1523	0.4554
	(0.4128)	(0.2982)	(0.3175)	(0.4214)	(0.2787)
being fidgety	0.1867	-0.0096	-0.4041	-0.3065	-0.2029
	(0.3950)	(0.2955)	(0.3155)	(0.4129)	(0.2763)
worrying about many things	0.7085*	0.6435**	0.3755	-0.7574*	0.3002
	(0.3933)	(0.2929)	(0.3088)	(0.4110)	(0.2748)
being irritable	0.3350	-0.1265	0.3459	-0.5732	0.2494
	(0.3832)	(0.2927)	(0.3047)	(0.4049)	(0.2690)
being upset by new situation	-0.9366**	-0.4386	0.0705	-0.8762*	-1.0929***
	(0.4280)	(0.3063)	(0.3314)	(0.4501)	(0.2888)
biting nails	-0.1887	-0.3727	0.1367	-0.0295	-0.4245
	(0.4701)	(0.3578)	(0.3658)	(0.4878)	(0.3319)
being disobedient	-0.1826	-0.5986**	-0.4912	-0.0479	-0.1582
	(0.4003)	(0.3000)	(0.3134)	(0.4207)	(0.2860)
Constant	28.3454***	35.6462***	35.2207***	33.7885***	28.4181***
	(1.2904)	(0.9289)	(1.0067)	(1.3213)	(0.8508)
Observations	1,618	1,618	1,618	1,618	1,618
R-squared	0.0402	0.0406	0.0292	0.0341	0.2333
Joint test of instruments					
$F(12, 1596)$	2.26	1.76	1.91	1.91	2.11
P-value	0.0077	0.0489	0.0288	0.0289	0.0139

Notes: Heteroskedasticity-robust standard errors in parentheses. *** Significant at the 1% level, ** Significant at the 5% level, * Significant at 10% level.

Table 1.9: Probit and IVprobit Estimates for Employment Status Model (No Self-employed Workers)

Variables	Probit (b)	Probit (b')	IVprobit (b)	IVprobit (b')
The Big Five				
extraversion	0.0006	0.0049	-0.0074	-0.0564
	(0.0010)		(0.0050)	
agreeableness	-0.0027**	-0.0152	0.0079*	0.0450
	(0.0013)		(0.0048)	
conscientiousness	0.0037***	0.0222	0.0134**	0.0800
	(0.0012)		(0.0056)	
emotional stability	0.0035***	0.0282	0.0020	0.0158
	(0.0009)		(0.0047)	
imagination	-0.0054***	-0.0315	-0.0396***	-0.2296
	(0.0015)		(0.0063)	
Other Explanatory Variables				
married	0.0808***	0.0364	0.0528**	0.0238
	(0.0132)		(0.0222)	
urban	-0.0098	-0.0043	-0.0041	-0.0018
	(0.0158)		(0.0187)	
wales	-0.0087	-0.0020	0.0301	0.0069
	(0.0287)		(0.0369)	
scotland	-0.0174	-0.0058	-0.0379*	-0.0126
	(0.0191)		(0.0223)	
compulsory education	0.0409**	0.0204	0.0433*	0.0216
	(0.0168)		(0.0222)	
extended secondary education	0.1211***	0.0359	0.1785***	0.0529
	(0.0338)		(0.0383)	
tertiary education	0.0905***	0.0398	0.1769***	0.0778
	(0.0244)		(0.0337)	
reading score at age 11	0.0043***	0.0256	0.0137***	0.0817
	(0.0014)		(0.0021)	
father's years of schooling	0.0103**	0.0223	0.0116**	0.0251
	(0.0052)		(0.0055)	
Observations	1,688		1,688	
Overidentification test of all instruments				
			$\chi^2(7) = 4.826$	
P-value			0.6811	

Notes: Marginal effects are reported rather than probit coefficients. Heteroskedasticity-robust standard errors in parentheses. *** Significant at the 1% level, ** Significant at the 5% level, * Significant at the 10% level. The Amemiya-Lee-Newey statistics for overidentification test of instruments are reported. The coefficient b' represents the change in employment probability for a one standard deviation increase in the independent variable ($b'_x = b_x * \sigma_x$).

Table 1.10: IVprobit Estimates - First-stage Regressions of Employments Status Model
(No Self-employed Workers)

Variables	extraversion	agreeableness	conscientiousness	emotional stability	imagination
married	1.0736***	-0.0095	0.8789***	1.4401***	-0.6749**
	(0.4145)	(0.3007)	(0.3388)	(0.4487)	(0.2843)
urban	0.5668	0.4830	-0.1030	-0.6888	0.1248
	(0.4149)	(0.3187)	(0.3307)	(0.4312)	(0.2968)
wales	1.3030*	0.2174	0.1223	1.6471**	0.8919
	(0.7743)	(0.6356)	(0.6733)	(0.7856)	(0.5532)
scotland	-0.5121	0.9313**	-0.4857	0.9449*	-0.4485
	(0.4981)	(0.4031)	(0.4153)	(0.5718)	(0.3563)
compulsory education	1.5175***	0.9862**	1.1520**	1.7726***	0.5138
	(0.5310)	(0.4185)	(0.4529)	(0.5927)	(0.3878)
extended secondary education	1.8316**	0.7015	1.9188***	2.8148***	2.3653***
	(0.7629)	(0.5830)	(0.6129)	(0.8482)	(0.5246)
tertiary education	2.2835***	1.6390***	1.4491***	2.8800***	3.1225***
	(0.6840)	(0.5097)	(0.5608)	(0.7182)	(0.5017)
reading score at age 11	-0.0045	0.0331	-0.0440	0.0059	0.2841***
	(0.0364)	(0.0275)	(0.0306)	(0.0391)	(0.0257)
father's years of schooling	0.0595	0.0593	0.0788	0.0691	0.0678
	(0.0911)	(0.0636)	(0.0630)	(0.0917)	(0.0571)
Instruments					
At Age 11					
hostility	0.1445	-0.1521*	-0.1099	-0.3458***	0.0899
	(0.1032)	(0.0873)	(0.0895)	(0.1261)	(0.0714)
immaturity	-0.4719***	-0.2591*	-0.4337***	0.0933	-0.0983
	(0.1792)	(0.1556)	(0.1680)	(0.2099)	(0.1284)
nervous symptoms	0.5105	-0.4835	0.1047	0.0391	-0.0252
	(0.4325)	(0.3843)	(0.4361)	(0.5240)	(0.3466)
At Age 7					
having difficulty concentrating	0.4347	0.2703	0.0200	0.1965	-0.0967
	(0.3969)	(0.3181)	(0.3303)	(0.4436)	(0.2515)
preferring to do things alone	-0.5864	-0.0442	0.3990	-0.8656**	0.1812
	(0.3761)	(0.2857)	(0.3109)	(0.4202)	(0.2377)
being bullied	-0.6681*	0.0507	-0.0268	-0.6173	0.5492**
	(0.3902)	(0.2935)	(0.3203)	(0.4273)	(0.2312)
being fidgety	0.1283	0.1047	-0.3679	0.0478	-0.3848*
	(0.3816)	(0.2960)	(0.3244)	(0.4235)	(0.2276)
worrying about many things	0.9661***	0.9205***	0.5248*	-0.3710	0.4110*
	(0.3680)	(0.2912)	(0.3032)	(0.4154)	(0.2306)
being irritable	0.3551	-0.1006	0.7565**	-0.3495	0.3231
	(0.3635)	(0.2884)	(0.3004)	(0.4038)	(0.2265)
being upset by new situation	-1.0486**	-0.5845*	0.1303	-0.8377*	-0.9425***
	(0.4151)	(0.3230)	(0.3325)	(0.4606)	(0.2752)
biting nails	-0.3744	-0.6430*	0.3445	-0.1107	-0.3283
	(0.4406)	(0.3522)	(0.3639)	(0.4981)	(0.2770)
being disobedient	-0.3415	-0.6151**	-0.4444	-0.2858	-0.1906
	(0.3676)	(0.2958)	(0.3121)	(0.4208)	(0.2311)
Constant	28.7910***	36.2649***	35.1249***	31.4992***	29.0238***
	(1.2563)	(0.9165)	(0.9621)	(1.3047)	(0.8009)
Observations	1,688	1,688	1,688	1,688	1,688
Joint test of instruments					
$\chi^2(12)$	29.21	32.45	28.22	26.94	28.42
P-value	0.0037	0.0012	0.0051	0.0079	0.0048

Note: Heteroskedasticity-robust standard errors in parentheses. *** Significant at the 1% level, ** Significant at the 5% level, * Significant at 10% level.

Table 1.11: MNL Results for Occupational Choice Model (No Self-employed Workers)

Variables	Managerial (b)	b'	Non-manual (b)	b'	Professional (b)	b'	Manual (b)	b'
The Big Five								
extraversion	0.0071***	0.0542	-0.0012	-0.0089	-0.0015	-0.0111	-0.0045***	-0.0343
	(0.0017)		(0.0017)		(0.0010)		(0.0017)	
agreeableness	-0.0059***	-0.0332	0.0055**	0.0313	0.0015	0.0084	-0.0012	-0.0065
	(0.0022)		(0.0024)		(0.0015)		(0.0022)	
conscientiousness	0.0040*	0.0232	-0.0035	-0.0207	0.0034***	0.0199	-0.0038*	-0.0224
	(0.0020)		(0.0022)		(0.0013)		(0.0020)	
emotional stability	0.0004	0.0033	-0.0015	-0.0119	0.0002	0.0018	0.0009	0.0068
	(0.0015)		(0.0016)		(0.0009)		(0.0015)	
imagination	0.0015	0.0087	-0.0008	-0.0045	0.0002	0.0012	-0.0009	-0.0053
	(0.0024)		(0.0025)		(0.0014)		(0.0024)	
Other Explanatory Variables								
married	0.1550***	0.0676	-0.0733***	-0.0320	0.0057	0.0025	-0.0874***	-0.0381
	(0.0278)		(0.0259)		(0.0159)		(0.0241)	
urban	-0.0163	-0.0072	0.0628**	0.0278	-0.0255*	-0.0113	-0.0210	-0.0093
	(0.0245)		(0.0264)		(0.0144)		(0.0246)	
wales	-0.1273**	-0.0290	0.0681	0.0155	-0.0383	-0.0087	0.0975**	0.0222
	(0.0580)		(0.0493)		(0.0355)		(0.0480)	
scotland	-0.0093	-0.0031	-0.0585	-0.0193	-0.0274	-0.0090	0.0951***	0.0314
	(0.0346)		(0.0378)		(0.0244)		(0.0305)	
compulsory education	0.0570	0.0285	0.0598	0.0299	-0.0204	-0.0102	-0.0963***	-0.0481
	(0.0423)		(0.0432)		(0.0293)		(0.0293)	
extended secondary education	0.1746***	0.0535	0.1598***	0.0490	-0.0362	-0.0111	-0.2982***	-0.0914
	(0.0500)		(0.0520)		(0.0369)		(0.0457)	
tertiary education	0.1322***	0.0591	0.1801***	0.0805	0.0607**	0.0271	-0.3729***	-0.1667
	(0.0462)		(0.0461)		(0.0283)		(0.0394)	
reading score at age 11	0.0036	0.0211	0.0082***	0.0482	0.0030**	0.0176	-0.0148***	-0.0869
	(0.0023)		(0.0023)		(0.0014)		(0.0022)	
father's years of schooling	0.0109**	0.0241	-0.0041	-0.0091	0.0060**	0.0132	-0.0128*	-0.0282
	(0.0050)		(0.0058)		(0.0023)		(0.0067)	
Observations	1,515		1,515		1,515		1,515	

Note: Marginal effects are reported rather than multinomial logit coefficients. Heteroskedasticity-robust standard errors in parentheses. *** Significant at the 1% level, ** Significant at the 5% level, * Significant at the 10% level. The coefficient b' represents the change in the probability of being in each occupation for a one standard deviation increase in the independent variable ($b'_x = b_x * \sigma_x$). The detailed information about each occupation group can be found in Appendix Table A.1.1.

Table 1.12: Two-stage MNL Estimates - First-stage Regressions of Occupational Choice Model (No Self-employed Workers)

Variables	extraversion	agreeableness	conscientiousness	emotional stability	imagination
married	1.0374**	0.1006	0.5685	0.8784*	-0.3826
	(0.4578)	(0.3194)	(0.3584)	(0.4675)	(0.3001)
urban	0.5572	0.5796*	0.0512	-0.5653	0.2510
	(0.4377)	(0.3309)	(0.3482)	(0.4474)	(0.3102)
wales	1.2856	0.3617	0.1934	1.6500**	1.1960**
	(0.8635)	(0.6962)	(0.7223)	(0.8303)	(0.5929)
scotland	-0.7854	1.1072***	-0.6554	0.5377	-0.6521*
	(0.5324)	(0.4066)	(0.4247)	(0.6080)	(0.3622)
compulsory education	1.4041**	1.1111**	1.3162***	1.5394**	0.8292**
	(0.5763)	(0.4602)	(0.5019)	(0.6252)	(0.4167)
extended secondary education	1.9410**	0.8878	2.0836***	2.7378***	2.7646***
	(0.7959)	(0.6233)	(0.6566)	(0.8657)	(0.5507)
tertiary education	2.2015***	1.7932***	1.6888***	2.5960***	3.5458***
	(0.7311)	(0.5480)	(0.5932)	(0.7379)	(0.5262)
reading score at age 11	-0.0218	0.0262	-0.0575*	-0.0102	0.2685***
	(0.0385)	(0.0293)	(0.0323)	(0.0398)	(0.0273)
father's years of schooling	0.0822	0.0900	0.0270	0.0441	0.0734
	(0.0946)	(0.0649)	(0.0639)	(0.0926)	(0.0582)
Instruments					
At Age 11					
hostility	0.1751	-0.1834*	-0.1472	-0.2824**	0.0167
	(0.1206)	(0.0981)	(0.1028)	(0.1419)	(0.0877)
immaturity	-0.3938*	-0.1533	-0.4747***	0.1263	-0.1080
	(0.2109)	(0.1607)	(0.1815)	(0.2210)	(0.1725)
nervous symptoms	0.5813	-0.7699*	-0.1160	-0.2658	-0.1719
	(0.5187)	(0.4061)	(0.4598)	(0.5458)	(0.3949)
At Age 7					
having difficulty concentrating	0.2001	0.1093	-0.0382	0.0553	-0.2821
	(0.4415)	(0.3321)	(0.3406)	(0.4504)	(0.3037)
preferring to do things alone	-0.5000	-0.1776	0.2791	-0.9705**	0.3458
	(0.4291)	(0.3020)	(0.3216)	(0.4227)	(0.2890)
being bullied	-0.7044	-0.0192	0.1537	-0.2313	0.3954
	(0.4329)	(0.3127)	(0.3327)	(0.4386)	(0.2897)
being fidgety	0.3623	0.0342	-0.4558	-0.0055	-0.1273
	(0.4148)	(0.3093)	(0.3277)	(0.4300)	(0.2866)
worrying about many things	0.8591**	1.0508***	0.4072	-0.7383*	0.3291
	(0.4118)	(0.3058)	(0.3149)	(0.4278)	(0.2841)
being irritable	0.4152	-0.1343	0.4756	-0.5742	0.4597
	(0.4029)	(0.3048)	(0.3165)	(0.4205)	(0.2793)
being upset by new situation	-1.1872***	-0.5898*	-0.0077	-0.9134*	-1.1230***
	(0.4465)	(0.3271)	(0.3423)	(0.4712)	(0.3045)
biting nails	-0.2829	-0.5248	0.3871	0.1548	-0.4513
	(0.4941)	(0.3701)	(0.3826)	(0.5092)	(0.3372)
being disobedient	-0.5448	-0.6648**	-0.3174	-0.2194	-0.2153
	(0.4146)	(0.3134)	(0.3258)	(0.4364)	(0.2941)
Constant	29.0684***	35.8413***	36.2910***	33.1183***	28.5404***
	(1.3555)	(0.9799)	(1.0206)	(1.3591)	(0.8805)
Observations	1,515	1,515	1,515	1,515	1,515
R-squared	0.0307	0.0474	0.0289	0.0400	0.2200
Joint test of instruments					
$F(12, 1493)$	1.92	2.72	1.78	2.22	2.04
P-value	0.0286	0.0012	0.0460	0.0090	0.0183

Notes: Heteroskedasticity-robust standard errors in parentheses. *** Significant at 1% level, ** Significant at 5% level, * Significant at 10% level.

Table 1.13: Two-stage MNL Results for Occupational Choice Model (No Self-employed Workers)

Variables	Managerial (b)	b'	Non-manual (b)	b'	Professional (b)	b'	Manual (b)	b'
The Big Five								
extraversion	0.0328*	0.2493	-0.0326*	-0.2473	-0.0038	-0.0288	0.0035	0.0268
	(0.0177)		(0.0170)		(0.0104)		(0.0142)	
agreeableness	-0.0034	-0.0192	0.0246	0.1392	-0.0204	-0.1154	-0.0008	-0.0046
	(0.0219)		(0.0217)		(0.0126)		(0.0184)	
conscientiousness	0.0385*	0.2251	-0.0401*	-0.2343	0.0116	0.0679	-0.0100	-0.0587
	(0.0225)		(0.0212)		(0.0134)		(0.0186)	
emotional stability	-0.0040	-0.0310	-0.0162	-0.1256	0.0018	0.0144	0.0183	0.1423
	(0.0134)		(0.0133)		(0.0077)		(0.0119)	
imagination	-0.0385	-0.2221	0.0436*	0.2514	0.0102	0.0589	-0.0153	-0.0882
	(0.0237)		(0.0227)		(0.0134)		(0.0208)	
Other Explanatory Variables								
married	0.0917*	0.0400	0.0146	0.0064	0.0120	0.0052	-0.1183***	-0.0516
	(0.0469)		(0.0466)		(0.0291)		(0.0398)	
urban	-0.0253	-0.0112	0.0512	0.0227	-0.0136	-0.0060	-0.0124	-0.0055
	(0.0345)		(0.0371)		(0.0195)		(0.0309)	
wales	-0.1188	-0.0271	0.0883	0.0201	-0.0455	-0.0104	0.0760	0.0173
	(0.0880)		(0.0877)		(0.1305)		(0.0715)	
scotland	0.0032	0.0010	-0.0838	-0.0276	0.0044	0.0014	0.0763	0.0251
	(0.0579)		(0.0608)		(0.0371)		(0.0466)	
compulsory education	0.0068	0.0034	0.1226*	0.0612	-0.0172	-0.0086	-0.1122**	-0.0561
	(0.0662)		(0.0655)		(0.0423)		(0.0496)	
extended secondary education	0.1664	0.0510	0.2052**	0.0629	-0.0649	-0.0199	-0.3068***	-0.0941
	(0.1066)		(0.1026)		(0.0712)		(0.0899)	
tertiary education	0.1649	0.0738	0.1590	0.0711	0.0521	0.0233	-0.3760***	-0.1681
	(0.1058)		(0.1003)		(0.0573)		(0.0844)	
reading score at age 11	0.0163**	0.0958	-0.0066	-0.0385	0.0015	0.0086	-0.0112*	-0.0659
	(0.0071)		(0.0070)		(0.0043)		(0.0063)	
father's years of schooling	0.0108*	0.0238	-0.0044	-0.0096	0.0076**	0.0167	-0.0140*	-0.0309
	(0.0065)		(0.0076)		(0.0035)		(0.0079)	
Observations	1,515		1,515		1,515		1,515	

Notes: Marginal effects are reported rather than multinomial logit coefficients. Bootstrapped standard errors in parentheses. *** Significant at the 1% level, ** Significant at 5% level, * Significant at the 10% level. The coefficient b' represents the change in the probability of being in each occupation for a one standard deviation increase in the independent variable ($b'_x = b_x * \sigma_x$). The detailed information about each occupation group can be found in Appendix Table A.1.1.

Table 1.14: OLS and 2SLS Regression Results for Earnings Model (No Self-employed Workers)

Variables	OLS (b)	OLS (b')	2SLS (b)	2SLS (b')
The Big Five				
extraversion	0.0096***	0.0727	0.0545*	0.4134
	(0.0026)		(0.0280)	
agreeableness	-0.0081***	-0.0455	0.0085	0.0482
	(0.0030)		(0.0380)	
conscientiousness	0.0105***	0.0618	0.0607*	0.3556
	(0.0036)		(0.0325)	
emotional stability	0.0051**	0.0394	0.0163	0.1265
	(0.0024)		(0.0204)	
imagination	0.0055	0.0321	-0.0374	-0.2167
	(0.0034)		(0.0334)	
Other Explanatory Variables				
married	0.1822***	0.0794	0.0726	0.0316
	(0.0376)		(0.0643)	
urban	-0.0468	-0.0207	-0.0887	-0.0392
	(0.0387)		(0.0561)	
wales	-0.2470***	-0.0570	-0.3187***	-0.0736
	(0.0616)		(0.1197)	
scotland	-0.0546	-0.0181	-0.0461	-0.0153
	(0.0433)		(0.0939)	
compulsory education	0.0447	0.0223	-0.1118	-0.0559
	(0.0468)		(0.0997)	
extended secondary education	0.2064**	0.0638	0.0569	0.0176
	(0.0953)		(0.1811)	
tertiary education	0.3080***	0.1388	0.2096	0.0944
	(0.0594)		(0.1538)	
reading score at age 11	0.0128***	0.0758	0.0268***	0.1587
	(0.0029)		(0.0098)	
father's years of schooling	0.0290***	0.0652	0.0223**	0.0501
	(0.0073)		(0.0099)	
Constant	4.9784***		2.2677	
	(0.2779)		(1.4198)	
Observations	1,330		1,330	
R-squared	0.1901		-	
Overidentification test of all instruments				
Hansen J-statistic			4.267	
P-value			0.7486	

Notes: Heteroskedasticity-robust standard errors in parentheses. *** Significant at the 1% level, ** Significant at the 5% level, * Significant at 10% level. R^2 is not reported for the 2SLS specification since it is negative. The coefficient b' represents the percentage change in earnings for a one standard deviation increase in the independent variable ($b'_x = b_x * \sigma_x$).

Table 1.15: 2SLS Estimates - First-stage Regressions of Earnings Model
(No Self-employed Workers)

Variables	extraversion	agreeableness	conscientiousness	emotional stability	imagination
married	1.2037**	0.2058	0.2887	0.6893	-0.3304
	(0.4857)	(0.3404)	(0.3839)	(0.5002)	(0.3160)
urban	0.8835*	0.6307*	0.2824	-0.2692	0.3975
	(0.4728)	(0.3568)	(0.3753)	(0.4769)	(0.3309)
wales	1.4523	0.9174	0.4975	2.3768***	1.5451**
	(0.8969)	(0.7132)	(0.7488)	(0.8543)	(0.6159)
scotland	-0.8314	1.1505***	-0.5159	0.4620	-0.5900
	(0.5689)	(0.4236)	(0.4520)	(0.6521)	(0.3838)
compulsory education	1.2759**	1.0283**	1.6373***	1.7827***	0.7692*
	(0.6177)	(0.4826)	(0.5434)	(0.6663)	(0.4497)
extended secondary education	2.0919**	0.9390	2.1654***	3.1439***	2.8189***
	(0.8647)	(0.6658)	(0.6894)	(0.9243)	(0.5938)
tertiary education	1.9899**	1.7227***	1.8266***	2.5769***	3.4842***
	(0.7812)	(0.5737)	(0.6362)	(0.7843)	(0.5615)
reading score at age 11	-0.0224	0.0121	-0.0557	-0.0105	0.2679***
	(0.0409)	(0.0309)	(0.0342)	(0.0426)	(0.0290)
father's years of schooling	0.0769	0.1117	0.0581	0.0375	0.0713
	(0.1014)	(0.0700)	(0.0699)	(0.0997)	(0.0624)
Instruments					
At Age 11					
hostility	0.2407*	-0.1081	-0.2261**	-0.1872	0.0013
	(0.1415)	(0.1050)	(0.1098)	(0.1646)	(0.1020)
immaturity	-0.5155**	-0.3024*	-0.4652**	0.1292	-0.1829
	(0.2219)	(0.1703)	(0.1910)	(0.2308)	(0.1833)
nervous symptoms	0.4718	-0.4214	-0.0319	-0.8309	-0.0835
	(0.5987)	(0.4736)	(0.5341)	(0.6031)	(0.4535)
At Age 7					
having difficulty concentrating	0.0963	-0.0352	-0.0340	-0.1067	-0.4630
	(0.4698)	(0.3542)	(0.3618)	(0.4780)	(0.3212)
preferring to do things alone	-0.3787	-0.0629	0.2788	-0.6812	0.3974
	(0.4558)	(0.3221)	(0.3438)	(0.4489)	(0.3066)
being bullied	-0.8033*	0.1271	0.2584	-0.3057	0.5192*
	(0.4575)	(0.3310)	(0.3512)	(0.4677)	(0.3041)
being fidgety	0.3555	0.0229	-0.5262	-0.0657	-0.1368
	(0.4483)	(0.3319)	(0.3508)	(0.4639)	(0.3062)
worrying about many things	0.9527**	1.1720***	0.2976	-0.7071	0.4187
	(0.4413)	(0.3238)	(0.3380)	(0.4577)	(0.3049)
being irritable	0.2081	-0.2901	0.4339	-0.7055	0.2882
	(0.4290)	(0.3240)	(0.3377)	(0.4460)	(0.2965)
being upset by new situation	-1.2207**	-0.6844**	0.0343	-0.9297*	-1.3239***
	(0.4767)	(0.3460)	(0.3682)	(0.5062)	(0.3198)
biting nails	-0.4631	-0.6121	0.4603	0.0495	-0.4740
	(0.5243)	(0.4017)	(0.4085)	(0.5484)	(0.3671)
being disobedient	-0.5243	-0.7867**	-0.4591	-0.2416	-0.3159
	(0.4436)	(0.3334)	(0.3444)	(0.4679)	(0.3171)
Constant	28.8298***	35.7874***	35.8517***	32.7623***	28.5701***
	(1.4445)	(1.0415)	(1.0961)	(1.4316)	(0.9316)
Observations	1,330	1,330	1,330	1,330	1,330
R-squared	0.0353	0.0500	0.0320	0.0396	0.2297
Joint test of instruments					
$F(12, 1308)$	1.97	2.86	1.84	1.90	2.55
P-value	0.0232	0.0007	0.0379	0.0304	0.0024

Notes: Heteroskedasticity-robust standard errors in parentheses. *** Significant at the 1% level, ** Significant at the 5% level, * Significant at 10% level.

Table 2.1a: Summary Statistics of Care Variables by Gender

	Female		Male	
Variables	Mean	S.D.	Mean	S.D.
care hours≥100	0.2231	0.4163	0.1676	0.3735
care hours≥200	0.1547	0.3616	0.1029	0.3038
care hours=0	0.7615	0.4262	0.7859	0.4102
0<care hours<200	0.0994	0.2992	0.1233	0.3288
200≤care hours<500	0.0611	0.2396	0.0497	0.2173
care hours≥500	0.0651	0.2467	0.0377	0.1904
care hours missing	0.0129	0.1127	0.0035	0.0587

Note: Based on different care definitions, the sample sizes differ. For females, the person-wave observations for "care hours≥200" is 14,139, and for all other care variables are 14,229; for males, the person-wave observations for "care hours≥200" are 10,389, and for all the other care variables are 10,405.

Table 2.1b: Summary Statistics by Gender and Care Status

	Female				Male			
	Caregiver		Non-caregiver		Caregiver		Non-caregiver	
Variables	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Dependent Variables								
could adjust work hours	0.4669	0.4990	0.4736	0.4993	0.4599	0.4985	0.4749	0.4994
flexibility index	0.3173	0.0792	0.3110	0.0819	0.2901	0.1036	0.2890	0.1053
top 3 flexible occupation categories	0.4605	0.4985	0.4172	0.4931	0.4112	0.4922	0.4205	0.4937
least flexible occupation category	0.0556	0.2291	0.0656	0.2476	0.1098	0.3127	0.1195	0.3244
Explanatory variables								
age	55.2621	5.0615	55.7592	5.3780	56.7150	4.4913	57.3026	4.4700
age squared	3,079.5142	541.7045	3,138.0096	576.5123	3,236.7540	488.2465	3,303.5686	487.9773
has spouse	0.7385	0.4395	0.7242	0.4469	0.8710	0.3353	0.8624	0.3445
has employed spouse	0.5306	0.4991	0.5168	0.4997	0.6095	0.4880	0.5737	0.4946
has spouse with ADL limitation	0.0580	0.2337	0.0627	0.2424	0.0797	0.2709	0.0704	0.2559
has kid younger than 18	0.1090	0.3117	0.1114	0.3146	0.1365	0.3434	0.1517	0.3588
household size	2.5375	1.1740	2.5043	1.2908	2.6026	1.1615	2.6440	1.3051
excellent health	0.1853	0.3886	0.1771	0.3818	0.1766	0.3814	0.1793	0.3836
very good health	0.3793	0.4853	0.3796	0.4853	0.3870	0.4872	0.3542	0.4783
good health	0.3106	0.4628	0.3076	0.4615	0.3108	0.4629	0.3259	0.4688
fair health	0.1059	0.3077	0.1210	0.3262	0.1101	0.3131	0.1193	0.3241
poor health	0.0189	0.1362	0.0147	0.1202	0.0155	0.1235	0.0212	0.1442
northeast	0.1711	0.3766	0.1682	0.3741	0.1542	0.3613	0.1611	0.3676
midwest	0.2710	0.4445	0.2529	0.4347	0.3022	0.4593	0.2569	0.4369
south	0.4175	0.4932	0.4030	0.4905	0.3888	0.4876	0.3854	0.4867
west and other regions	0.1405	0.3476	0.1758	0.3807	0.1548	0.3618	0.1966	0.3975
homeowner	0.8771	0.3283	0.8467	0.3603	0.8813	0.3235	0.8641	0.3427
household non-housing wealth/10^3	207.5875	578.9181	181.4931	545.2247	203.0526	467.9357	210.6427	1654.6520
non-labor income/10^3	8.5045	22.2966	8.2584	52.3568	11.6358	96.8673	8.7485	49.1522
experience	31.1972	9.1164	30.6549	9.7652	37.4455	7.1057	37.5174	7.8935
experience squared	1,056.3497	530.3165	1,035.0743	559.0967	1,452.6302	475.5879	1,469.8536	511.8805
Person-wave observations	3,174		11,055		1,744		8,661	

Note: The summary statistics are computed based on the sample used in the baseline model specification, i.e., the sample with the dependent variable as "could adjust work hours" and the care variable as "care hours$\geq$100".

Table 2.2: FE Model of Care's Effect on Job Choice - Flexible Schedule

Variables	Female I	Female II	Female III	Male I	Male II	Male III
care hours≥100	0.0341**			-0.0232		
	(0.0133)			(0.0162)		
care hours≥200		0.0145			-0.0115	
		(0.0147)			(0.0194)	
0<care hours<200			0.0296**			-0.0104
			(0.0151)			(0.0172)
200≤care hours<500			0.0152			-0.0040
			(0.0195)			(0.0243)
care hours≥500			0.0039			-0.0028
			(0.0212)			(0.0293)
care hours missing			-0.0230			-0.0327
			(0.0210)			(0.0256)
age	-0.0509**	-0.0489**	-0.0633***	-0.0125	-0.0122	-0.0110
	(0.0216)	(0.0217)	(0.0198)	(0.0299)	(0.0300)	(0.0285)
age squared	0.0004***	0.0004**	0.0005***	0.0003	0.0003	0.0002
	(0.0002)	(0.0002)	(0.0001)	(0.0002)	(0.0002)	(0.0002)
has spouse	0.0091	0.0076	0.0032	-0.0961**	-0.0973**	-0.0807**
	(0.0286)	(0.0289)	(0.0252)	(0.0391)	(0.0393)	(0.0351)
has employed spouse	0.0185	0.0192	0.0138	0.0243	0.0242	0.0172
	(0.0161)	(0.0162)	(0.0154)	(0.0178)	(0.0178)	(0.0172)
has spouse with ADL limitation	0.0238	0.0233	0.0225	-0.0217	-0.0229	-0.0181
	(0.0215)	(0.0216)	(0.0209)	(0.0253)	(0.0254)	(0.0245)
has kid younger than 18	0.0233	0.0228	0.0295	-0.0079	-0.0073	-0.0107
	(0.0216)	(0.0217)	(0.0199)	(0.0232)	(0.0232)	(0.0220)
household size	-0.0005	-0.0006	0.0011	0.0062	0.0062	0.0082
	(0.0059)	(0.0060)	(0.0056)	(0.0078)	(0.0078)	(0.0075)
excellent health	-0.0436	-0.0451	-0.0654*	0.0762	0.0762	0.0572
	(0.0412)	(0.0417)	(0.0394)	(0.0496)	(0.0497)	(0.0469)
very good health	-0.0260	-0.0265	-0.0418	0.0780	0.0772	0.0611
	(0.0393)	(0.0398)	(0.0377)	(0.0476)	(0.0477)	(0.0449)
good health	-0.0310	-0.0317	-0.0512	0.0645	0.0634	0.0455
	(0.0381)	(0.0386)	(0.0368)	(0.0466)	(0.0467)	(0.0439)
fair health	-0.0112	-0.0123	-0.0228	0.0720	0.0725	0.0431
	(0.0369)	(0.0374)	(0.0357)	(0.0454)	(0.0456)	(0.0426)
midwest	-0.0094	0.0203	0.0186	0.1143	0.0901	0.1609
	(0.1006)	(0.1003)	(0.0949)	(0.1293)	(0.1287)	(0.1252)
south	0.0907	0.1085	0.1281*	0.0734	0.0776	0.1086
	(0.0810)	(0.0802)	(0.0752)	(0.0993)	(0.0992)	(0.0984)
west and other regions	0.0112	0.0273	0.0937	0.1512	0.1466	0.2104*
	(0.1012)	(0.1009)	(0.0957)	(0.1267)	(0.1265)	(0.1216)
homeowner	-0.0142	-0.0153	-0.0278	-0.0302	-0.0293	-0.0348
	(0.0246)	(0.0246)	(0.0226)	(0.0305)	(0.0305)	(0.0295)
household non-housing wealth/10^6	0.0224**	0.0223**	0.0254***	0.0012	0.0013	0.0014
	(0.0090)	(0.0091)	(0.0088)	(0.0015)	(0.0015)	(0.0015)
non-labor income/10^6	0.0004	0.0017	0.0167	0.1545**	0.1528**	0.1614**
	(0.0874)	(0.0882)	(0.0940)	(0.0656)	(0.0658)	(0.0705)
experience	-0.0405***	-0.0380***	-0.0338***	-0.0587***	-0.0590***	-0.0526**
	(0.0141)	(0.0142)	(0.0125)	(0.0218)	(0.0218)	(0.0207)
experience squared	0.0002**	0.0002**	0.0002**	-0.0001	-0.0001	-0.0001
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Observations	14,229	14,117	15,930	10,405	10,386	11,095
Number of Individuals	3,840	3,813	4,235	3,045	3,042	3,193

Notes: Time fixed effects are included. Robust standard errors in parentheses. *** Significant at 1% level, ** Significant at 5% level, * Significant at 10% level.

Table 2.3: FE Model of Care's Effect on Job Choice - Flexibility Index

Variables	Female I	Female II	Female III	Male I	Male II	Male III
care hours≥100	-0.0001			0.0032**		
	(0.0011)			(0.0016)		
care hours≥200		0.0000			0.0035*	
		(0.0011)			(0.0020)	
0<care hours<200			-0.0000			0.0012
			(0.0012)			(0.0016)
200≤care hours<500			-0.0002			0.0026
			(0.0014)			(0.0021)
care hours≥500			0.0009			0.0051
			(0.0018)			(0.0033)
care hours missing			-0.0019			0.0005
			(0.0019)			(0.0022)
age	-0.0006	-0.0004	0.0002	0.0003	0.0003	0.0011
	(0.0021)	(0.0021)	(0.0020)	(0.0042)	(0.0043)	(0.0038)
age squared	0.0000	0.0000	0.0000	-0.0000	-0.0000	-0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
has spouse	-0.0006	-0.0005	-0.0028	0.0005	0.0006	0.0015
	(0.0025)	(0.0025)	(0.0024)	(0.0042)	(0.0042)	(0.0039)
has employed spouse	-0.0006	-0.0006	-0.0004	0.0035**	0.0036**	0.0035**
	(0.0014)	(0.0014)	(0.0014)	(0.0018)	(0.0018)	(0.0017)
has spouse with ADL limitation	0.0007	0.0007	0.0011	-0.0011	-0.0011	0.0003
	(0.0016)	(0.0016)	(0.0015)	(0.0025)	(0.0025)	(0.0024)
has kid younger than 18	0.0026	0.0026	0.0030*	-0.0019	-0.0020	-0.0013
	(0.0019)	(0.0019)	(0.0017)	(0.0026)	(0.0026)	(0.0024)
household size	0.0002	0.0002	0.0004	0.0002	0.0002	0.0001
	(0.0006)	(0.0006)	(0.0005)	(0.0007)	(0.0007)	(0.0007)
excellent health	0.0050	0.0049	0.0055*	-0.0013	-0.0013	-0.0007
	(0.0032)	(0.0032)	(0.0029)	(0.0047)	(0.0047)	(0.0043)
very good health	0.0014	0.0013	0.0020	-0.0019	-0.0019	-0.0014
	(0.0029)	(0.0029)	(0.0027)	(0.0045)	(0.0044)	(0.0041)
good health	0.0011	0.0010	0.0011	0.0016	0.0016	0.0016
	(0.0028)	(0.0029)	(0.0026)	(0.0044)	(0.0044)	(0.0040)
fair health	0.0016	0.0015	0.0013	-0.0014	-0.0014	-0.0013
	(0.0028)	(0.0028)	(0.0026)	(0.0044)	(0.0044)	(0.0040)
midwest	-0.0017	-0.0006	-0.0004	0.0082	0.0094	0.0100
	(0.0163)	(0.0163)	(0.0150)	(0.0182)	(0.0185)	(0.0168)
south	-0.0111	-0.0113	-0.0079	0.0115	0.0113	0.0139
	(0.0135)	(0.0135)	(0.0125)	(0.0110)	(0.0110)	(0.0100)
west and other regions	-0.0102	-0.0097	-0.0041	0.0073	0.0073	0.0096
	(0.0135)	(0.0135)	(0.0129)	(0.0189)	(0.0190)	(0.0171)
homeowner	-0.0038	-0.0037	-0.0040	-0.0056*	-0.0056*	-0.0063**
	(0.0028)	(0.0028)	(0.0025)	(0.0033)	(0.0033)	(0.0032)
household non-housing wealth/10^6	-0.0005	-0.0005	-0.0005	0.0001	0.0001	0.0001
	(0.0006)	(0.0006)	(0.0006)	(0.0001)	(0.0001)	(0.0001)
non-labor income/10^6	-0.0042	-0.0043	-0.0035	0.0156	0.0156	0.0167
	(0.0033)	(0.0032)	(0.0034)	(0.0126)	(0.0126)	(0.0117)
experience	-0.0033*	-0.0030	-0.0029	0.0067**	0.0067**	0.0060**
	(0.0020)	(0.0020)	(0.0019)	(0.0029)	(0.0029)	(0.0028)
experience squared	0.0000	0.0000	0.0000	-0.0000	-0.0000	-0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Observations	14,621	14,507	16,371	10,948	10,924	11,649
Number of Individuals	3,849	3,822	4,239	3,103	3,099	3,245

Notes: Time fixed effects are included. Robust standard errors in parentheses. *** Significant at 1% level, ** Significant at 5% level, * Significant at 10% level. Regressors are jointly insignificant for the models Female I, II, III.

Table 2.4: FE Model of Care's Effect on Job Choice - Top 3 Flexible Occupation Categories

Variables	Female I	Female II	Female III	Male I	Male II	Male III
care hours≥100	-0.0037			0.0152**		
	(0.0063)			(0.0076)		
care hours≥200		-0.0018			0.0129	
		(0.0070)			(0.0093)	
0<care hours<200			-0.0044			0.0057
			(0.0074)			(0.0071)
200≤care hours<500			-0.0104			0.0055
			(0.0086)			(0.0099)
care hours≥500			0.0149			0.0282*
			(0.0106)			(0.0171)
care hours missing			-0.0133			0.0028
			(0.0106)			(0.0103)
age	0.0046	0.0049	0.0058	0.0087	0.0082	0.0076
	(0.0129)	(0.0129)	(0.0126)	(0.0187)	(0.0189)	(0.0165)
age squared	0.0000	0.0000	-0.0000	-0.0001	-0.0001	-0.0001
	(0.0001)	(0.0001)	(0.0001)	(0.0002)	(0.0002)	(0.0001)
has spouse	-0.0003	-0.0003	-0.0155	-0.0056	-0.0053	-0.0005
	(0.0150)	(0.0151)	(0.0140)	(0.0196)	(0.0197)	(0.0176)
has employed spouse	-0.0050	-0.0052	-0.0025	0.0122	0.0122	0.0116
	(0.0081)	(0.0082)	(0.0080)	(0.0083)	(0.0083)	(0.0079)
has spouse with ADL limitation	0.0073	0.0067	0.0103	-0.0021	-0.0015	0.0039
	(0.0093)	(0.0093)	(0.0090)	(0.0112)	(0.0113)	(0.0105)
has kid younger than 18	0.0099	0.0097	0.0120	-0.0137	-0.0138	-0.0110
	(0.0104)	(0.0104)	(0.0100)	(0.0111)	(0.0112)	(0.0108)
household size	0.0022	0.0023	0.0025	-0.0001	-0.0001	-0.0001
	(0.0034)	(0.0035)	(0.0031)	(0.0032)	(0.0032)	(0.0035)
excellent health	0.0003	-0.0002	0.0037	0.0069	0.0067	0.0139
	(0.0185)	(0.0186)	(0.0166)	(0.0182)	(0.0182)	(0.0166)
very good health	-0.0174	-0.0184	-0.0135	0.0033	0.0033	0.0079
	(0.0169)	(0.0171)	(0.0153)	(0.0170)	(0.0170)	(0.0154)
good health	-0.0161	-0.0165	-0.0140	0.0123	0.0123	0.0163
	(0.0163)	(0.0165)	(0.0147)	(0.0166)	(0.0166)	(0.0149)
fair health	-0.0115	-0.0117	-0.0122	0.0049	0.0048	0.0075
	(0.0168)	(0.0170)	(0.0151)	(0.0166)	(0.0166)	(0.0151)
midwest	-0.0019	-0.0024	0.0125	-0.0114	-0.0040	-0.0036
	(0.1038)	(0.1047)	(0.1005)	(0.0817)	(0.0826)	(0.0751)
south	-0.0362	-0.0363	-0.0150	-0.0097	-0.0102	0.0004
	(0.0825)	(0.0825)	(0.0831)	(0.0541)	(0.0541)	(0.0498)
west and other regions	-0.0105	-0.0097	0.0304	0.0243	0.0254	0.0373
	(0.0905)	(0.0906)	(0.0899)	(0.1040)	(0.1042)	(0.0932)
homeowner	-0.0119	-0.0107	-0.0105	-0.0212	-0.0216	-0.0239
	(0.0158)	(0.0157)	(0.0142)	(0.0160)	(0.0160)	(0.0151)
household non-housing wealth/10^6	-0.0008	-0.0010	-0.0010	0.0000	-0.0000	0.0001
	(0.0031)	(0.0031)	(0.0031)	(0.0003)	(0.0003)	(0.0003)
non-labor income/10^6	-0.0496**	-0.0506**	-0.0457**	-0.0249	-0.0241	-0.0194
	(0.0235)	(0.0235)	(0.0226)	(0.0163)	(0.0163)	(0.0150)
experience	-0.0188	-0.0179	-0.0158	0.0372***	0.0368***	0.0393***
	(0.0122)	(0.0122)	(0.0116)	(0.0137)	(0.0137)	(0.0129)
experience squared	0.0001	0.0001	0.0001	-0.0000	-0.0000	-0.0000
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Observations	14,621	14,507	16,371	10,948	10,924	11,649
Number of Individuals	3,849	3,822	4,239	3,103	3,099	3,245

Notes: Time fixed effects are included. Robust standard errors in parentheses. *** Significant at 1% level, ** Significant at 5% level, * Significant at 10% level. Regressors are jointly insignificant for the models Female I, II, III.

Table 2.5: FE Model of Care's Effect on Job Choice - Least Flexible Occupation Category

Variables	Female I	Female II	Female III	Male I	Male II	Male III
care hours≥100	0.0008			-0.0048		
	(0.0033)			(0.0051)		
care hours≥200		0.0003			-0.0098	
		(0.0033)			(0.0066)	
0<care hours<200			-0.0002			-0.0049
			(0.0038)			(0.0045)
200≤care hours<500			-0.0052			-0.0117*
			(0.0044)			(0.0068)
care hours≥500			0.0028			-0.0153
			(0.0050)			(0.0109)
care hours missing			0.0006			-0.0047
			(0.0062)			(0.0075)
age	0.0043	0.0039	0.0055	0.0109	0.0098	0.0080
	(0.0054)	(0.0054)	(0.0049)	(0.0114)	(0.0114)	(0.0116)
age squared	-0.0001**	-0.0001**	-0.0001**	-0.0000	-0.0000	-0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0001)	(0.0001)	(0.0001)
has spouse	0.0066	0.0059	0.0100	0.0112	0.0110	0.0194*
	(0.0078)	(0.0079)	(0.0071)	(0.0105)	(0.0105)	(0.0109)
has employed spouse	-0.0033	-0.0034	-0.0027	-0.0054	-0.0055	-0.0074
	(0.0043)	(0.0043)	(0.0041)	(0.0053)	(0.0053)	(0.0054)
has spouse with ADL limitation	-0.0014	-0.0013	-0.0015	0.0079	0.0080	0.0032
	(0.0043)	(0.0044)	(0.0039)	(0.0096)	(0.0096)	(0.0099)
has kid younger than 18	-0.0026	-0.0025	-0.0032	-0.0109	-0.0097	-0.0108
	(0.0058)	(0.0059)	(0.0051)	(0.0081)	(0.0080)	(0.0076)
household size	-0.0013	-0.0014	-0.0012	-0.0031	-0.0034	-0.0029
	(0.0015)	(0.0015)	(0.0014)	(0.0025)	(0.0025)	(0.0024)
excellent health	-0.0246**	-0.0246**	-0.0244***	-0.0037	-0.0039	-0.0035
	(0.0105)	(0.0106)	(0.0093)	(0.0181)	(0.0181)	(0.0177)
very good health	-0.0174*	-0.0173*	-0.0179**	0.0023	0.0017	0.0008
	(0.0097)	(0.0098)	(0.0085)	(0.0177)	(0.0177)	(0.0174)
good health	-0.0134	-0.0135	-0.0136	-0.0116	-0.0118	-0.0118
	(0.0095)	(0.0096)	(0.0083)	(0.0175)	(0.0176)	(0.0172)
fair health	-0.0186**	-0.0187**	-0.0152**	0.0018	0.0019	0.0011
	(0.0089)	(0.0089)	(0.0077)	(0.0179)	(0.0180)	(0.0176)
midwest	0.0030	0.0037	0.0017	0.0314	0.0336	0.0179
	(0.0396)	(0.0396)	(0.0345)	(0.0553)	(0.0560)	(0.0484)
south	0.0052	0.0051	0.0024	-0.0137	-0.0130	-0.0248
	(0.0407)	(0.0407)	(0.0349)	(0.0418)	(0.0419)	(0.0358)
west and other regions	-0.0109	-0.0108	-0.0099	0.0004	0.0018	-0.0203
	(0.0365)	(0.0365)	(0.0318)	(0.0700)	(0.0702)	(0.0566)
homeowner	0.0102	0.0104	0.0109	0.0172*	0.0173*	0.0156*
	(0.0076)	(0.0077)	(0.0069)	(0.0097)	(0.0097)	(0.0094)
household non-housing wealth/10^6	0.0010	0.0010	0.0010	-0.0004	-0.0004	-0.0003
	(0.0011)	(0.0011)	(0.0011)	(0.0003)	(0.0003)	(0.0003)
non-labor income/10^6	-0.0056	-0.0055	-0.0077	-0.0684*	-0.0677*	-0.0503**
	(0.0064)	(0.0064)	(0.0070)	(0.0350)	(0.0347)	(0.0228)
experience	0.0077*	0.0073	0.0069	-0.0028	-0.0028	-0.0001
	(0.0046)	(0.0045)	(0.0042)	(0.0090)	(0.0090)	(0.0082)
experience squared	-0.0000	-0.0000	-0.0000	0.0000	0.0000	0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0001)	(0.0001)	(0.0000)
Observations	14,621	14,507	16,371	10,948	10,924	11,649
Number of Individuals	3,849	3,822	4,239	3,103	3,099	3,245

Notes: Time fixed effects are included. Robust standard errors in parentheses. *** Significant at 1% level, ** Significant at 5% level, * Significant at 10% level. Regressors are jointly insignificant for the model Male III.

Table 2.6: Pooled OLS and FE for Care's Effects on Job Choice - Baseline Models

	Flexible Schedule				Top 3 Flexible Occupation Categories			
	Female I		Male I		Female I		Male I	
Variables	Pooled OLS	FE	Pooled OLS	FE	Pooled OLS	FE	Pooled OLS	FE
care hours$\geq$100	0.0040	0.0341**	-0.0145	-0.0232	0.0153	-0.0037	-0.0232	0.0152**
	(0.0116)	(0.0133)	(0.0150)	(0.0162)	(0.0122)	(0.0063)	(0.0144)	(0.0076)
Observations	14,229		10,405		14,621		10,948	
Number of Individuals	3,840		3,045		3,849		3,103	

Notes: Time fixed effects are included. Robust standard errors in parentheses. *** Significant at 1% level, ** Significant at 5% level, * Significant at 10% level.

Table 2.7: FE Model of Care's Effect on Job Choice - Flexible Schedule (Selected Sample)

Variables	Female I	Female II	Female III	Male I	Male II	Male III
care hours≥100	0.0242			-0.0195		
	(0.0148)			(0.0177)		
care hours≥200		0.0067			-0.0034	
		(0.0162)			(0.0210)	
0<care hours<200			0.0417***			-0.0036
			(0.0159)			(0.0182)
200≤care hours<500			0.0209			0.0069
			(0.0204)			(0.0258)
care hours≥500			0.0267			-0.0082
			(0.0228)			(0.0303)
care hours missing			-0.0125			-0.0516
			(0.0281)			(0.0329)
age	-0.0487*	-0.0483*	-0.0560**	-0.0130	-0.0123	-0.0023
	(0.0260)	(0.0261)	(0.0247)	(0.0349)	(0.0351)	(0.0345)
age squared	0.0004**	0.0004**	0.0005***	0.0004	0.0003	0.0003
	(0.0002)	(0.0002)	(0.0002)	(0.0003)	(0.0003)	(0.0003)
has spouse	0.0319	0.0266	0.0187	-0.0870*	-0.0877*	-0.0873*
	(0.0368)	(0.0373)	(0.0336)	(0.0519)	(0.0526)	(0.0469)
has employed spouse	0.0246	0.0264	0.0254	0.0463**	0.0465**	0.0398*
	(0.0196)	(0.0197)	(0.0191)	(0.0214)	(0.0215)	(0.0208)
has spouse with ADL limitation	0.0292	0.0288	0.0306	-0.0445	-0.0464	-0.0445
	(0.0257)	(0.0258)	(0.0254)	(0.0298)	(0.0301)	(0.0293)
has kid younger than 18	0.0426*	0.0422*	0.0483**	0.0271	0.0276	0.0248
	(0.0252)	(0.0253)	(0.0235)	(0.0277)	(0.0277)	(0.0264)
household size	0.0018	0.0017	0.0045	0.0086	0.0085	0.0093
	(0.0082)	(0.0083)	(0.0079)	(0.0097)	(0.0097)	(0.0093)
excellent health	-0.0711	-0.0718	-0.0952*	0.1002*	0.1000*	0.0943*
	(0.0561)	(0.0570)	(0.0539)	(0.0564)	(0.0566)	(0.0535)
very good health	-0.0553	-0.0561	-0.0779	0.0976*	0.0963*	0.0966*
	(0.0537)	(0.0547)	(0.0518)	(0.0538)	(0.0541)	(0.0509)
good health	-0.0347	-0.0354	-0.0648	0.0823	0.0804	0.0813
	(0.0521)	(0.0531)	(0.0503)	(0.0530)	(0.0532)	(0.0502)
fair health	-0.0162	-0.0184	-0.0392	0.0745	0.0753	0.0652
	(0.0522)	(0.0531)	(0.0504)	(0.0520)	(0.0524)	(0.0489)
midwest	0.1397	0.1901*	0.1659	-0.0644	-0.1036	-0.0419
	(0.1143)	(0.1114)	(0.1089)	(0.1517)	(0.1520)	(0.1476)
south	0.2365**	0.2687***	0.2558***	-0.0343	-0.0277	-0.0064
	(0.0958)	(0.0927)	(0.0902)	(0.1055)	(0.1061)	(0.1061)
west and other regions	0.1679	0.1972*	0.2414**	0.0525	0.0440	0.0929
	(0.1127)	(0.1118)	(0.1052)	(0.1380)	(0.1383)	(0.1379)
homeowner	-0.0065	-0.0079	-0.0098	-0.0268	-0.0255	-0.0190
	(0.0324)	(0.0323)	(0.0303)	(0.0377)	(0.0377)	(0.0373)
household non-housing wealth/10^6	0.0214**	0.0208*	0.0222**	0.0012	0.0014	0.0018
	(0.0109)	(0.0110)	(0.0108)	(0.0028)	(0.0029)	(0.0030)
non-labor income/10^6	0.1429	0.1455	0.1389	0.1691**	0.1659**	0.1800**
	(0.1216)	(0.1236)	(0.1192)	(0.0811)	(0.0820)	(0.0891)
experience	-0.0482***	-0.0434**	-0.0399**	-0.0489*	-0.0497*	-0.0352
	(0.0175)	(0.0176)	(0.0162)	(0.0258)	(0.0259)	(0.0247)
experience squared	0.0002*	0.0002	0.0001	-0.0002	-0.0002	-0.0003
	(0.0001)	(0.0001)	(0.0001)	(0.0002)	(0.0002)	(0.0002)
Observations	9,024	8,916	9,729	7,171	7,154	7,574
Number of Individuals	2,560	2,531	2,729	2,197	2,195	2,291

Notes: Time fixed effects are included. Robust standard errors in parentheses. *** Significant at 1% level, ** Significant at 5% level, * Significant at 10% level. The selected sample includes those individuals aged between 25 and 64 who are employed for at least two waves of the survey and have at least one alive parent/in-law in the current wave.

Table 2.8: FE Model of Care's Effect on Job Choice - Flexibility Index (Selected Sample)

Variables	Female I	Female II	Female III	Male I	Male II	Male III
care hours≥100	-0.0003			0.0023		
	(0.0012)			(0.0016)		
care hours≥200		-0.0004			0.0027	
		(0.0012)			(0.0020)	
0<care hours<200			0.0006			0.0006
			(0.0013)			(0.0016)
200≤care hours<500			0.0003			0.0018
			(0.0014)			(0.0022)
care hours≥500			0.0008			0.0038
			(0.0018)			(0.0036)
care hours missing			0.0000			0.0012
			(0.0025)			(0.0028)
age	0.0015	0.0018	0.0011	0.0049	0.0050	0.0058
	(0.0025)	(0.0025)	(0.0025)	(0.0051)	(0.0051)	(0.0046)
age squared	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
has spouse	-0.0015	-0.0010	-0.0035	0.0009	0.0011	0.0019
	(0.0032)	(0.0033)	(0.0031)	(0.0055)	(0.0055)	(0.0056)
has employed spouse	0.0005	0.0006	0.0010	0.0035	0.0035	0.0035
	(0.0016)	(0.0016)	(0.0016)	(0.0023)	(0.0023)	(0.0022)
has spouse with ADL limitation	0.0025	0.0024	0.0025	-0.0011	-0.0011	-0.0003
	(0.0020)	(0.0020)	(0.0020)	(0.0029)	(0.0029)	(0.0029)
has kid younger than 18	0.0041*	0.0041*	0.0047**	-0.0023	-0.0024	-0.0012
	(0.0021)	(0.0021)	(0.0020)	(0.0029)	(0.0029)	(0.0028)
household size	0.0003	0.0004	0.0005	-0.0001	-0.0000	-0.0003
	(0.0006)	(0.0007)	(0.0006)	(0.0008)	(0.0008)	(0.0009)
excellent health	0.0045	0.0044	0.0047	-0.0035	-0.0036	-0.0031
	(0.0039)	(0.0039)	(0.0037)	(0.0060)	(0.0060)	(0.0057)
very good health	0.0034	0.0032	0.0033	-0.0026	-0.0026	-0.0023
	(0.0036)	(0.0037)	(0.0035)	(0.0058)	(0.0059)	(0.0055)
good health	0.0041	0.0040	0.0040	-0.0005	-0.0005	-0.0005
	(0.0035)	(0.0036)	(0.0034)	(0.0058)	(0.0058)	(0.0054)
fair health	0.0051	0.0051	0.0044	-0.0035	-0.0036	-0.0035
	(0.0034)	(0.0035)	(0.0033)	(0.0060)	(0.0060)	(0.0056)
midwest	-0.0044	-0.0024	-0.0012	0.0192	0.0216	0.0176
	(0.0222)	(0.0223)	(0.0208)	(0.0188)	(0.0192)	(0.0179)
south	-0.0177	-0.0178	-0.0146	0.0055	0.0050	0.0045
	(0.0189)	(0.0188)	(0.0178)	(0.0111)	(0.0112)	(0.0109)
west and other regions	-0.0114	-0.0107	-0.0058	-0.0028	-0.0024	-0.0002
	(0.0176)	(0.0176)	(0.0171)	(0.0118)	(0.0120)	(0.0114)
homeowner	-0.0033	-0.0030	-0.0034	-0.0094**	-0.0094**	-0.0093**
	(0.0035)	(0.0035)	(0.0032)	(0.0042)	(0.0042)	(0.0041)
household non-housing wealth/10^6	-0.0007	-0.0007	-0.0007	-0.0001	-0.0001	-0.0001
	(0.0008)	(0.0008)	(0.0008)	(0.0001)	(0.0001)	(0.0001)
non-labor income/10^6	0.0005	0.0001	0.0006	0.0219*	0.0219*	0.0228**
	(0.0048)	(0.0048)	(0.0049)	(0.0118)	(0.0117)	(0.0102)
experience	-0.0031	-0.0024	-0.0037	0.0028	0.0027	0.0026
	(0.0025)	(0.0025)	(0.0024)	(0.0031)	(0.0031)	(0.0030)
experience squared	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Observations	9,277	9,169	10,013	7,540	7,519	7,945
Number of Individuals	2,575	2,547	2,747	2,244	2,241	2,332

Notes: Time fixed effects are included. Robust standard errors in parentheses. *** Significant at 1% level, ** Significant at 5% level, * Significant at 10% level. Regressors are jointly insignificant for the models Female I, II, III, and Male I, II, III. The selected sample includes those individuals aged between 25 and 64 who are employed for at least two waves of the survey and have at least one alive parent/in-law in the current wave.

Table 2.9: FE Model of Care's Effect on Job Choice - Top 3 Flexible Occupation Categories (Selected Sample)

Variables	Female I	Female II	Female III	Male I	Male II	Male III
care hours≥100	-0.0053			0.0148*		
	(0.0069)			(0.0077)		
care hours≥200		-0.0056			0.0136	
		(0.0076)			(0.0092)	
0<care hours<200			0.0001			0.0059
			(0.0075)			(0.0076)
200≤care hours<500			-0.0058			0.0062
			(0.0091)			(0.0106)
care hours≥500			0.0165			0.0239
			(0.0112)			(0.0176)
care hours missing			-0.0110			0.0073
			(0.0140)			(0.0135)
age	0.0114	0.0124	0.0094	0.0373*	0.0365*	0.0332*
	(0.0147)	(0.0147)	(0.0152)	(0.0214)	(0.0214)	(0.0192)
age squared	-0.0000	-0.0000	-0.0000	-0.0003	-0.0003	-0.0002
	(0.0001)	(0.0001)	(0.0001)	(0.0002)	(0.0002)	(0.0002)
has spouse	-0.0226	-0.0224	-0.0317*	-0.0058	-0.0055	0.0098
	(0.0191)	(0.0192)	(0.0186)	(0.0263)	(0.0265)	(0.0255)
has employed spouse	0.0040	0.0037	0.0075	0.0123	0.0119	0.0118
	(0.0090)	(0.0090)	(0.0090)	(0.0107)	(0.0106)	(0.0103)
has spouse with ADL limitation	0.0130	0.0119	0.0156	0.0011	0.0017	0.0081
	(0.0119)	(0.0119)	(0.0117)	(0.0135)	(0.0137)	(0.0131)
has kid younger than 18	0.0106	0.0100	0.0146	-0.0164	-0.0167	-0.0124
	(0.0118)	(0.0118)	(0.0114)	(0.0129)	(0.0129)	(0.0124)
household size	0.0054	0.0057	0.0050	-0.0014	-0.0014	-0.0024
	(0.0036)	(0.0037)	(0.0033)	(0.0036)	(0.0036)	(0.0042)
excellent health	0.0082	0.0074	0.0116	-0.0058	-0.0063	0.0028
	(0.0211)	(0.0215)	(0.0203)	(0.0226)	(0.0226)	(0.0208)
very good health	0.0009	-0.0003	0.0021	-0.0038	-0.0039	0.0036
	(0.0196)	(0.0200)	(0.0189)	(0.0218)	(0.0218)	(0.0200)
good health	0.0065	0.0063	0.0092	-0.0008	-0.0009	0.0064
	(0.0188)	(0.0192)	(0.0182)	(0.0213)	(0.0213)	(0.0194)
fair health	0.0163	0.0165	0.0117	-0.0050	-0.0053	-0.0007
	(0.0191)	(0.0195)	(0.0187)	(0.0226)	(0.0226)	(0.0208)
midwest	0.0777	0.0778	0.0916	-0.0032	0.0103	-0.0109
	(0.1354)	(0.1370)	(0.1344)	(0.0882)	(0.0890)	(0.0850)
south	0.0220	0.0223	0.0302	-0.0579	-0.0592	-0.0656
	(0.1049)	(0.1048)	(0.1098)	(0.0674)	(0.0673)	(0.0657)
west and other regions	0.0822	0.0833	0.1122	-0.0572	-0.0542	-0.0446
	(0.1033)	(0.1032)	(0.1072)	(0.0604)	(0.0604)	(0.0602)
homeowner	-0.0102	-0.0073	-0.0120	-0.0503**	-0.0506**	-0.0483**
	(0.0190)	(0.0188)	(0.0176)	(0.0199)	(0.0199)	(0.0191)
household non-housing wealth/10^6	-0.0026	-0.0027	-0.0026	-0.0001	-0.0002	-0.0001
	(0.0038)	(0.0038)	(0.0038)	(0.0004)	(0.0004)	(0.0004)
non-labor income/10^6	-0.0138	-0.0167	-0.0138	-0.0295	-0.0283	-0.0207
	(0.0275)	(0.0279)	(0.0278)	(0.0193)	(0.0191)	(0.0173)
experience	-0.0165	-0.0145	-0.0249	0.0085	0.0079	0.0166
	(0.0163)	(0.0165)	(0.0157)	(0.0152)	(0.0150)	(0.0145)
experience squared	0.0001	0.0001	0.0000	0.0001	0.0001	0.0001
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Observations	9,277	9,169	10,013	7,540	7,519	7,945
Number of Individuals	2,575	2,547	2,747	2,244	2,241	2,332

Notes: Time fixed effects are included. Robust standard errors in parentheses. *** Significant at 1% level, ** Significant at 5% level, * Significant at 10% level. Regressors are jointly insignificant for the models Female I, II, III and Male II, III. The selected sample includes those individuals aged between 25 and 64 who are employed for at least two waves of the survey and have at least one alive parent/in-law in the current wave.

Table 2.10: FE Model of Care's Effect on Job Choice - Least Flexible Occupation Category (Selected Sample)

Variables	Female I	Female II	Female III	Male I	Male II	Male III
care hours≥100	-0.0011			-0.0063		
	(0.0035)			(0.0053)		
care hours≥200		-0.0023			-0.0082	
		(0.0034)			(0.0071)	
0<care hours<200			-0.0003			-0.0053
			(0.0040)			(0.0046)
200≤care hours<500			-0.0058			-0.0082
			(0.0041)			(0.0071)
care hours≥500			0.0047			-0.0152
			(0.0051)			(0.0122)
care hours missing			-0.0083			-0.0012
			(0.0082)			(0.0096)
age	-0.0012	-0.0020	0.0031	0.0156	0.0122	0.0087
	(0.0066)	(0.0067)	(0.0063)	(0.0142)	(0.0138)	(0.0136)
age squared	-0.0000	-0.0000	-0.0000	-0.0001	-0.0000	-0.0000
	(0.0001)	(0.0001)	(0.0000)	(0.0001)	(0.0001)	(0.0001)
has spouse	-0.0075	-0.0095	-0.0002	0.0128	0.0126	0.0322*
	(0.0096)	(0.0097)	(0.0098)	(0.0132)	(0.0133)	(0.0166)
has employed spouse	-0.0039	-0.0041	-0.0030	-0.0045	-0.0048	-0.0065
	(0.0046)	(0.0047)	(0.0046)	(0.0063)	(0.0063)	(0.0063)
has spouse with ADL limitation	-0.0038	-0.0037	-0.0034	0.0187*	0.0189*	0.0175
	(0.0044)	(0.0044)	(0.0039)	(0.0106)	(0.0107)	(0.0111)
has kid younger than 18	-0.0089	-0.0087	-0.0096*	-0.0058	-0.0039	-0.0089
	(0.0057)	(0.0057)	(0.0051)	(0.0088)	(0.0086)	(0.0083)
household size	0.0007	0.0005	0.0003	-0.0033	-0.0037	-0.0028
	(0.0021)	(0.0021)	(0.0018)	(0.0030)	(0.0030)	(0.0029)
excellent health	-0.0237*	-0.0240*	-0.0218	0.0095	0.0098	0.0081
	(0.0142)	(0.0145)	(0.0133)	(0.0237)	(0.0238)	(0.0241)
very good health	-0.0202	-0.0204	-0.0191	0.0101	0.0097	0.0085
	(0.0137)	(0.0139)	(0.0127)	(0.0234)	(0.0235)	(0.0239)
good health	-0.0192	-0.0196	-0.0179	-0.0049	-0.0048	-0.0056
	(0.0136)	(0.0139)	(0.0126)	(0.0232)	(0.0233)	(0.0237)
fair health	-0.0270**	-0.0276**	-0.0250**	0.0072	0.0072	0.0049
	(0.0126)	(0.0128)	(0.0115)	(0.0243)	(0.0244)	(0.0246)
midwest	-0.0167	-0.0167	-0.0136	0.0270	0.0283	0.0188
	(0.0641)	(0.0640)	(0.0574)	(0.0329)	(0.0343)	(0.0293)
south	0.0054	0.0050	0.0048	0.0040	0.0050	-0.0006
	(0.0675)	(0.0675)	(0.0598)	(0.0182)	(0.0185)	(0.0153)
west and other regions	-0.0301	-0.0300	-0.0235	0.0589	0.0602	0.0422
	(0.0596)	(0.0596)	(0.0534)	(0.0392)	(0.0395)	(0.0312)
homeowner	0.0161	0.0165	0.0147	0.0060	0.0061	0.0058
	(0.0106)	(0.0107)	(0.0099)	(0.0087)	(0.0087)	(0.0082)
household non-housing wealth/10^6	0.0005	0.0006	0.0005	0.0000	0.0000	-0.0000
	(0.0011)	(0.0011)	(0.0011)	(0.0003)	(0.0003)	(0.0003)
non-labor income/10^6	-0.0081	-0.0078	-0.0110	-0.0855***	-0.0850***	-0.0586***
	(0.0127)	(0.0125)	(0.0136)	(0.0295)	(0.0293)	(0.0206)
experience	0.0058	0.0049	0.0046	-0.0043	-0.0043	-0.0039
	(0.0063)	(0.0061)	(0.0058)	(0.0087)	(0.0088)	(0.0081)
experience squared	-0.0000	-0.0000	-0.0000	0.0000	0.0000	0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0001)	(0.0001)	(0.0001)
Observations	9,277	9,169	10,013	7,540	7,519	7,945
Number of Individuals	2,575	2,547	2,747	2,244	2,241	2,332

Notes: Time fixed effects are included. Robust standard errors in parentheses. *** Significant at 1% level, ** Significant at 5% level, * Significant at 10% level. Regressors are jointly insignificant for the models Female I, II, III. The selected sample includes those individuals aged between 25 and 64 who are employed for at least two waves of the survey and have at least one alive parent/in-law in the current wave.

Table 3.1: Summary Statistics

	Urban (N=3235)				Rural (N=666)			
Variables	Mean	S.D.	Min	Max	Mean	S.D.	Min	Max
log of monthly income	5.910	0.993	2.303	10.003	5.178	0.831	2.708	9.297
health index	2.368	0.693	1.000	3.000	2.224	0.702	1.000	3.000
excellent health	0.491	0.500	0.000	1.000	0.383	0.486	0.000	1.000
good health	0.385	0.487	0.000	1.000	0.458	0.499	0.000	1.000
poor health	0.124	0.329	0.000	1.000	0.159	0.366	0.000	1.000
has health problem	0.194	0.396	0.000	1.000	0.194	0.395	0.000	1.000
years of schooling	8.325	4.132	0.000	21.000	4.856	3.581	0.000	17.000
starting age	15.001	4.376	5.000	31.000	12.285	4.337	5.000	31.000
age	32.511	9.585	18.000	55.000	30.937	9.241	18.000	55.000
age squared	1148.805	662.261	324.000	3025.000	1042.369	628.206	324.000	3025.000
male	0.571	0.495	0.000	1.000	0.697	0.460	0.000	1.000
white	0.486	0.500	0.000	1.000	0.434	0.496	0.000	1.000
black	0.068	0.252	0.000	1.000	0.071	0.256	0.000	1.000
other race	0.446	0.497	0.000	1.000	0.495	0.500	0.000	1.000
current GDP/capita	5.908	2.871	1.670	10.649	5.469	2.745	1.670	10.649
father schooling missing	0.289	0.453	0.000	1.000	0.458	0.499	0.000	1.000
father illiterate	0.029	0.168	0.000	1.000	0.063	0.243	0.000	1.000
father lower primary	0.466	0.499	0.000	1.000	0.414	0.493	0.000	1.000
father upper primary	0.098	0.297	0.000	1.000	0.029	0.167	0.000	1.000
father highschool	0.072	0.258	0.000	1.000	0.023	0.148	0.000	1.000
father college	0.047	0.212	0.000	1.000	0.014	0.116	0.000	1.000
mother schooling missing	0.324	0.468	0.000	1.000	0.514	0.500	0.000	1.000
mother illiterate	0.030	0.171	0.000	1.000	0.062	0.241	0.000	1.000
mother lower primary	0.437	0.496	0.000	1.000	0.353	0.478	0.000	1.000
mother upper primary	0.099	0.298	0.000	1.000	0.036	0.187	0.000	1.000
mother highschool	0.086	0.280	0.000	1.000	0.030	0.171	0.000	1.000
mother college	0.024	0.153	0.000	1.000	0.006	0.077	0.000	1.000
father not work	0.114	0.318	0.000	1.000	0.059	0.235	0.000	1.000
father employee	0.508	0.500	0.000	1.000	0.535	0.499	0.000	1.000
father self-employed	0.274	0.446	0.000	1.000	0.357	0.480	0.000	1.000
father employer	0.054	0.227	0.000	1.000	0.020	0.138	0.000	1.000
father occupation missing	0.050	0.218	0.000	1.000	0.030	0.171	0.000	1.000
mother not work	0.564	0.496	0.000	1.000	0.506	0.500	0.000	1.000
mother employee	0.249	0.433	0.000	1.000	0.224	0.417	0.000	1.000
mother self-employed	0.122	0.327	0.000	1.000	0.117	0.322	0.000	1.000
mother employer	0.010	0.099	0.000	1.000	0.002	0.039	0.000	1.000
mother unsalaried	0.043	0.204	0.000	1.000	0.135	0.342	0.000	1.000
mother occupation missing	0.011	0.106	0.000	1.000	0.017	0.128	0.000	1.000
age12_gdp/capita	4.099	3.204	0.333	12.983	4.042	2.967	0.440	12.983
age7_teacher/school	5.271	4.235	1.543	27.243	5.327	3.923	1.543	23.968
age11_teacher/school	6.060	4.636	1.566	27.390	6.088	4.339	1.566	23.968
age7_hospital/1000 residents	0.039	0.014	0.007	0.121	0.042	0.015	0.007	0.121
age7_bed/1000 residents	3.378	1.733	0.107	7.868	3.407	1.677	0.335	7.295
age7_doctor/1000 residents	0.809	0.616	0.050	3.006	0.796	0.565	0.050	2.934

Notes: There are only 2 individuals reporting father's occupation type as "unsalaried", thus I collapse the type "unsalaried" and "self-employed" for father's occupation.

Table 3.2: OLS of Income Model

Variables	Urban	Rural
starting age	-0.002	0.008
	(0.004)	(0.007)
years of schooling	0.097***	0.093***
	(0.004)	(0.012)
excellent health	0.189***	0.094
	(0.042)	(0.077)
good health	0.087**	0.120
	(0.042)	(0.084)
age	0.096***	0.074***
	(0.010)	(0.021)
age squared	-0.001***	-0.001***
	(0.000)	(0.000)
male	0.617***	0.662***
	(0.026)	(0.062)
black	-0.131***	-0.014
	(0.050)	(0.123)
other race	-0.066**	0.041
	(0.028)	(0.060)
current GDP/capita	0.061***	0.035***
	(0.005)	(0.011)
father schooling missing	-0.069*	0.014
	(0.036)	(0.071)
father illiterate	-0.071	-0.156
	(0.066)	(0.118)
father upper primary	0.021	0.151
	(0.048)	(0.135)
father highschool	0.193***	0.120
	(0.060)	(0.258)
father college	0.330***	0.639*
	(0.077)	(0.369)
mother schooling missing	0.008	-0.070
	(0.033)	(0.065)
mother illiterate	0.066	0.052
	(0.067)	(0.119)
mother upper primary	0.056	-0.121
	(0.045)	(0.140)
mother highschool	0.105*	-0.087
	(0.054)	(0.281)
mother college	0.041	-0.014
	(0.092)	(0.543)
Constant	2.227***	2.430***
	(0.179)	(0.386)
Observations	3,235	666
R-squared	0.479	0.318

Notes: Clustered robust standard errors are reported in parentheses.
*** Significant at 1% level** Significant at 5% level, * Significant at 10% level.

Table 3.3a: IV Estimates - First-stage Regression of Income Model for Urban

Variables	Starting age_u	Schooling_u	Excellent health_u	Good health_u
age	0.337***	0.505***	-0.012	0.021**
	(0.085)	(0.060)	(0.011)	(0.010)
age squared	-0.005***	-0.006***	0.000	-0.000*
	(0.001)	(0.001)	(0.000)	(0.000)
male	-1.986***	-0.877***	0.057***	-0.028*
	(0.152)	(0.116)	(0.017)	(0.017)
black	-0.619**	-1.128***	-0.053	0.054
	(0.248)	(0.240)	(0.036)	(0.038)
other race	-0.433***	-1.071***	-0.029	0.032
	(0.156)	(0.129)	(0.021)	(0.022)
current GDP/capita	-0.215***	-0.223***	0.008	-0.002
	(0.046)	(0.040)	(0.005)	(0.005)
father schooling missing	-1.025***	-1.183***	-0.075***	0.051**
	(0.202)	(0.174)	(0.023)	(0.024)
father illiterate	-1.254***	-2.232***	-0.052	0.049
	(0.455)	(0.390)	(0.051)	(0.052)
father upper primary	0.396	0.903***	0.021	-0.002
	(0.249)	(0.208)	(0.032)	(0.031)
father highschool	1.089***	1.528***	0.043	-0.031
	(0.272)	(0.251)	(0.039)	(0.036)
father college	1.695***	2.572***	0.077	-0.014
	(0.343)	(0.278)	(0.048)	(0.046)
mother schooling missing	-0.783***	-2.049***	-0.030	0.004
	(0.179)	(0.158)	(0.022)	(0.021)
mother illiterate	-1.136***	-1.596***	-0.103*	0.037
	(0.437)	(0.356)	(0.053)	(0.055)
mother upper primary	0.905***	1.040***	0.021	-0.006
	(0.231)	(0.208)	(0.028)	(0.029)
mother highschool	1.487***	1.962***	0.055	-0.017
	(0.288)	(0.198)	(0.035)	(0.033)
mother college	1.593***	2.955***	0.150**	-0.122**
	(0.515)	(0.388)	(0.060)	(0.060)
father occupation missing	-0.178	0.082	0.053	-0.053
	(0.343)	(0.287)	(0.043)	(0.042)
father not work	-0.710***	0.273	0.050*	-0.036
	(0.219)	(0.204)	(0.029)	(0.028)
father self-employed	-0.399**	0.016	0.020	-0.027
	(0.181)	(0.154)	(0.021)	(0.021)
father employer	-0.190	0.850***	0.019	-0.019
	(0.347)	(0.260)	(0.040)	(0.039)
mother occupation missing	0.617	0.599	0.131	-0.121
	(0.609)	(0.509)	(0.088)	(0.086)
mother not work	0.689***	0.649***	0.042**	-0.010
	(0.161)	(0.140)	(0.020)	(0.021)
mother self-employed	0.063	0.504**	0.066**	-0.063**
	(0.256)	(0.205)	(0.030)	(0.030)
mother employer	-0.348	0.479	0.161**	-0.070
	(0.803)	(0.543)	(0.081)	(0.073)
mother unsalaried	-1.651***	-0.453	0.033	-0.046
	(0.448)	(0.341)	(0.046)	(0.047)
age12_gdp/capita	0.079	0.284***	0.010	-0.012
	(0.063)	(0.048)	(0.008)	(0.007)
age7_teacher/school	-0.049	-0.042	-0.001	0.001
	(0.051)	(0.039)	(0.008)	(0.007)
age11_teacher/school	-0.017	-0.075*	0.003	-0.004
	(0.047)	(0.040)	(0.005)	(0.005)
age7_hospital/1000 residents	8.895	7.492	-1.301	2.300**
	(9.075)	(6.350)	(1.103)	(1.064)
age7_bed/1000 residents	-0.148	-0.053	0.026*	-0.016
	(0.138)	(0.107)	(0.015)	(0.015)
age7_doctor/1000 residents	1.267***	0.641**	-0.173***	0.149***
	(0.441)	(0.279)	(0.047)	(0.042)
Constant	11.146***	-0.075	0.792***	-0.066
	(1.660)	(1.133)	(0.205)	(0.194)
Observations	3,235	3,235	3,235	3,235
R-squared	0.208	0.378	0.058	0.023
Test of excluded	F(15, 457)	F(15, 457)	F(15, 457)	F(15, 457)
instruments	7.42	5.82	3.12	3.48
P-value	0.000	0.000	0.000	0.000

Notes: Clustered robust standard errors are reported in parentheses. *** Significant at 1% level, ** Significant at 5% level, * Significant at 10% level.

Table 3.3b: IV Estimates - First-stage Regression of Income Model for Rural

Variables	Starting age_r	Schooling_r	Excellent health_r	Good health_r
age	0.224	-0.157	0.005	-0.012
	(0.176)	(0.133)	(0.023)	(0.023)
age squared	-0.003	0.002	-0.000	0.000
	(0.002)	(0.002)	(0.000)	(0.000)
male	-2.439***	-0.989***	0.104***	-0.082*
	(0.359)	(0.259)	(0.040)	(0.046)
black	-0.898	-0.152	-0.035	-0.017
	(0.577)	(0.434)	(0.075)	(0.091)
other race	-0.607*	-0.337	-0.163***	0.106**
	(0.356)	(0.269)	(0.042)	(0.043)
current GDP/capita	-0.257	0.039	0.024	-0.009
	(0.157)	(0.124)	(0.019)	(0.020)
father schooling missing	-0.530	-0.934***	-0.133***	0.119***
	(0.367)	(0.285)	(0.044)	(0.044)
father illiterate	0.534	-1.446***	-0.211**	0.224***
	(0.752)	(0.451)	(0.084)	(0.082)
father upper primary	1.670**	1.750***	-0.041	-0.011
	(0.780)	(0.644)	(0.127)	(0.116)
father highschool	2.369**	3.574***	-0.131	0.195
	(1.167)	(0.579)	(0.107)	(0.142)
father college	3.361**	4.864***	0.127	0.033
	(1.537)	(1.055)	(0.170)	(0.189)
mother schooling missing	-1.111***	-1.907***	0.025	-0.057
	(0.385)	(0.276)	(0.048)	(0.050)
mother illiterate	-1.415**	-2.263***	0.088	-0.267***
	(0.686)	(0.489)	(0.097)	(0.092)
mother upper primary	2.003**	1.304*	0.016	-0.017
	(0.841)	(0.705)	(0.110)	(0.117)
mother highschool	2.777**	1.632**	0.192*	-0.291***
	(1.168)	(0.808)	(0.113)	(0.092)
mother college	-1.614	0.313	0.207	-0.101
	(2.441)	(1.665)	(0.245)	(0.260)
father occupation missing	-0.194	-0.287	0.095	-0.034
	(1.032)	(0.786)	(0.115)	(0.128)
father not work	-0.252	0.352	0.072	-0.072
	(0.574)	(0.472)	(0.076)	(0.080)
father self-employed	-0.030	0.380	-0.065	0.029
	(0.358)	(0.257)	(0.050)	(0.050)
father employer	-0.106	0.957	-0.025	-0.029
	(0.897)	(1.054)	(0.138)	(0.152)
mother occupation missing	-0.263	0.437	-0.043	-0.183
	(0.892)	(0.745)	(0.126)	(0.155)
mother not work	0.138	0.631**	0.090*	-0.082
	(0.427)	(0.311)	(0.050)	(0.053)
mother self-employed	0.285	0.922**	-0.036	-0.084
	(0.535)	(0.461)	(0.069)	(0.072)
mother employer	-0.196	-0.942**	-0.242***	0.392***
	(0.606)	(0.398)	(0.078)	(0.079)
mother unsalaried	-1.535**	-0.422	-0.048	0.002
	(0.607)	(0.372)	(0.066)	(0.072)
age12_gdp/capita	0.175	0.072	-0.017	0.003
	(0.144)	(0.129)	(0.020)	(0.022)
age7_teacher/school	0.123	-0.058	0.012	0.001
	(0.096)	(0.066)	(0.012)	(0.012)
age11_teacher/school	-0.010	0.029	-0.028***	0.023**
	(0.098)	(0.066)	(0.011)	(0.010)
age7_hospital/1000 residents	4.582	-37.066**	-1.813	0.446
	(17.658)	(15.121)	(2.274)	(2.220)
age7_bed/1000 residents	-0.446	-0.242	0.002	0.019
	(0.318)	(0.232)	(0.039)	(0.039)
age7_doctor/1000 residents	0.445	0.192	0.134	-0.207**
	(0.726)	(0.575)	(0.105)	(0.105)
Constant	12.076***	11.198***	0.469	0.713
	(3.384)	(2.655)	(0.441)	(0.438)
Observations	666	666	666	666
R-squared	0.276	0.413	0.117	0.065
Test of excluded	F(15, 269)	F(15, 269)	F(15, 269)	F(15, 269)
instruments	1.77	4.26	5.24	8.23
P-value	0.039	0.000	0.000	0.000

Notes: Clustered robust standard errors are reported in parentheses. *** Significant at 1% level ** Significant at 5% level, * Significant at 10% level.

Table 3.4: IV Estimates - Second-stage Regression of Income Model

Variables	Urban	Rural
starting age	0.018	0.167**
	(0.041)	(0.071)
years of schooling	0.083*	0.111*
	(0.049)	(0.064)
excellent health	1.640**	0.591
	(0.812)	(0.659)
good health	0.219	0.895
	(0.839)	(0.737)
age	0.084***	0.046
	(0.017)	(0.033)
age squared	-0.001**	-0.000
	(0.000)	(0.000)
male	0.567***	1.089***
	(0.078)	(0.183)
black	-0.033	0.169
	(0.082)	(0.151)
other race	-0.018	0.134
	(0.057)	(0.103)
current GDP/capita	0.051***	0.050**
	(0.010)	(0.020)
father schooling missing	0.025	0.118
	(0.065)	(0.124)
father illiterate	-0.004	-0.246
	(0.123)	(0.254)
father upper primary	0.010	-0.130
	(0.071)	(0.222)
father highschool	0.150	-0.476
	(0.105)	(0.387)
father college	0.231*	-0.074
	(0.126)	(0.544)
mother schooling missing	0.034	0.167
	(0.089)	(0.145)
mother illiterate	0.213*	0.478*
	(0.112)	(0.272)
mother upper primary	0.037	-0.490**
	(0.070)	(0.234)
mother highschool	0.045	-0.475
	(0.093)	(0.405)
mother college	-0.118	0.173
	(0.145)	(0.684)
Constant	1.325*	-0.391
	(0.698)	(1.046)
Observations	3,235	666
R-squared	0.033	-
Overidentification test of all instruments		
Hansen J-statistic	6.703	15.300
P-value	0.823	0.169

Notes: Clustered robust standard errors are reported in parentheses. *** Significant at 1% level, ** Significant at 5% level, * Significant at 10% level. R^2 is not reported for the rural sample since it's negative.

Table 3.5: Ordered Probit Estimates of Health Model (3.2a)

Variables	Urban_1	Urban_2	Urban_3	Rural_1	Rural_2	Rural_3
starting age	-0.001	-0.002	0.003	-0.005*	-0.003*	0.008*
	(0.001)	(0.001)	(0.002)	(0.003)	(0.002)	(0.004)
years of schooling	-0.007***	-0.008***	0.015***	-0.005	-0.004	0.009
	(0.001)	(0.001)	(0.003)	(0.003)	(0.002)	(0.006)
age	0.006	0.007	-0.013	-0.000	-0.000	0.000
	(0.004)	(0.005)	(0.009)	(0.011)	(0.008)	(0.019)
age squared	-0.000	-0.000	0.000	0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
male	-0.038***	-0.041***	0.079***	-0.080***	-0.044***	0.124***
	(0.008)	(0.009)	(0.017)	(0.024)	(0.012)	(0.034)
black	0.010	0.010	-0.020	0.031	0.018	-0.049
	(0.016)	(0.016)	(0.031)	(0.051)	(0.023)	(0.074)
other race	0.003	0.003	-0.005	0.077***	0.053***	-0.131***
	(0.009)	(0.010)	(0.020)	(0.024)	(0.018)	(0.040)
current GDP/capita	-0.007***	-0.008***	0.015***	-0.004	-0.003	0.007
	(0.002)	(0.003)	(0.005)	(0.007)	(0.005)	(0.012)
father schooling missing	0.020*	0.021*	-0.040*	0.052**	0.035**	-0.087**
	(0.011)	(0.011)	(0.022)	(0.025)	(0.016)	(0.040)
father illiterate	0.002	0.002	-0.003	0.076	0.031***	-0.107
	(0.023)	(0.025)	(0.049)	(0.057)	(0.012)	(0.067)
father upper primary	-0.005	-0.005	0.010	0.041	0.021	-0.062
	(0.014)	(0.017)	(0.031)	(0.077)	(0.027)	(0.104)
father highschool	-0.006	-0.007	0.014	0.061	0.027	-0.088
	(0.018)	(0.021)	(0.039)	(0.088)	(0.021)	(0.108)
father college	-0.028	-0.036	0.064	-0.068	-0.079	0.147
	(0.019)	(0.029)	(0.047)	(0.060)	(0.111)	(0.171)
mother schooling missing	0.001	0.001	-0.002	-0.022	-0.015	0.037
	(0.011)	(0.012)	(0.023)	(0.027)	(0.019)	(0.046)
mother illiterate	0.046	0.039**	-0.085*	0.013	0.008	-0.021
	(0.029)	(0.019)	(0.048)	(0.055)	(0.032)	(0.088)
mother upper primary	-0.002	-0.002	0.004	-0.023	-0.019	0.042
	(0.013)	(0.015)	(0.028)	(0.051)	(0.049)	(0.101)
mother highschool	-0.014	-0.016	0.030	-0.020	-0.016	0.035
	(0.016)	(0.020)	(0.035)	(0.065)	(0.060)	(0.124)
mother college	-0.036	-0.051	0.087	-0.085	-0.116	0.200
	(0.023)	(0.041)	(0.064)	(0.075)	(0.181)	(0.255)
age7_hospital/1000 residents	0.389	0.432	-0.821	0.420	0.294	-0.714
	(0.422)	(0.472)	(0.894)	(1.197)	(0.836)	(2.031)
age7_bed/1000 residents	-0.015**	-0.016**	0.031**	-0.019	-0.013	0.032
	(0.006)	(0.007)	(0.013)	(0.021)	(0.014)	(0.035)
age7_doctor/1000 residents	0.068***	0.075***	-0.142***	0.020	0.014	-0.033
	(0.015)	(0.018)	(0.033)	(0.047)	(0.032)	(0.079)
Observations	3,235	3,235	3,235	666	666	666

Notes: Marginal effects rather than ordered probit estimates are reported. Clustered robust standard errors are reported in parentheses. *** Significant at 1% level, ** Significant at 5% level, * Significant at 10% level. Column 1, 2 and 3 correspond to the marginal effects on the probability that the health index equals 1, 2, and 3 for urban residents, column 4, 5 and 6 correspond to the marginal effects for rural residents.

Table 3.6: IV Ordered Probit Estimates - First-stage Regression of Health Model (3.2a)

Variables	starting age_u	schooling_u	starting age_r	schooling_r
age	0.341***	0.495***	0.235	-0.159
	(0.086)	(0.060)	(0.180)	(0.134)
age squared	-0.005***	-0.006***	-0.003	0.002
	(0.001)	(0.001)	(0.002)	(0.002)
male	-1.986***	-0.876***	-2.461***	-0.974***
	(0.151)	(0.115)	(0.350)	(0.252)
black	-0.619**	-1.127***	-0.840	-0.193
	(0.247)	(0.239)	(0.567)	(0.430)
other race	-0.433***	-1.070***	-0.597*	-0.343
	(0.156)	(0.129)	(0.348)	(0.265)
current GDP/capita	-0.216***	-0.219***	-0.199	-0.002
	(0.046)	(0.040)	(0.154)	(0.111)
father schooling missing	-1.021***	-1.194***	-0.551	-0.920***
	(0.202)	(0.172)	(0.363)	(0.276)
father illiterate	-1.255***	-2.229***	0.517	-1.433***
	(0.453)	(0.388)	(0.730)	(0.443)
father upper primary	0.396	0.904***	1.716**	1.722***
	(0.248)	(0.207)	(0.760)	(0.624)
father highschool	1.089***	1.529***	2.453**	3.512***
	(0.271)	(0.250)	(1.129)	(0.565)
father college	1.695***	2.572***	3.581**	4.708***
	(0.341)	(0.277)	(1.499)	(1.027)
mother schooling missing	-0.781***	-2.054***	-1.061***	-1.942***
	(0.177)	(0.156)	(0.377)	(0.271)
mother illiterate	-1.135***	-1.600***	-1.403**	-2.271***
	(0.434)	(0.355)	(0.671)	(0.473)
mother upper primary	0.903***	1.046***	1.938**	1.349*
	(0.230)	(0.207)	(0.806)	(0.690)
mother highschool	1.483***	1.973***	2.988**	1.482*
	(0.288)	(0.196)	(1.174)	(0.777)
mother college	1.586***	2.972***	-1.513	0.247
	(0.513)	(0.383)	(2.283)	(1.695)
age7_hospital/1000 residents	8.983	7.298	-1.957	-31.567**
	(9.036)	(6.296)	(18.744)	(14.792)
age7_bed/1000 residents	-0.149	-0.050	-0.368	-0.307
	(0.137)	(0.107)	(0.332)	(0.236)
age7_doctor/1000 residents	1.289***	0.581**	0.920	-0.131
	(0.446)	(0.287)	(0.709)	(0.543)
father occupation missing	-0.200	0.141	-0.106	-0.347
	(0.341)	(0.287)	(0.888)	(0.658)
father not work	-0.721***	0.304	0.159	0.065
	(0.219)	(0.193)	(0.519)	(0.418)
father self-employed	-0.400**	0.021	0.098	0.290
	(0.178)	(0.150)	(0.326)	(0.247)
father employer	-0.161	0.775***	0.459	0.563
	(0.354)	(0.263)	(0.980)	(0.916)
mother occupation missing	0.565	0.737	-0.153	0.363
	(0.608)	(0.518)	(0.923)	(0.637)
mother not work	0.673***	0.694***	0.557*	0.340
	(0.169)	(0.130)	(0.334)	(0.335)
mother self-employed	0.049	0.541***	0.538	0.748
	(0.254)	(0.194)	(0.508)	(0.478)
mother employer	-0.458	0.772	-0.743*	-0.554
	(0.823)	(0.558)	(0.447)	(0.431)
mother unsalaried	-1.659***	-0.430	-1.131*	-0.703**
	(0.444)	(0.336)	(0.671)	(0.335)
age12_gdp/capita	0.085	0.268***	0.101	0.125
	(0.066)	(0.053)	(0.147)	(0.095)
age7_teacher/school	-0.049	-0.041	0.056	-0.011
	(0.051)	(0.043)	(0.105)	(0.065)
age11_teacher/school	-0.022	-0.063	-0.023	0.039
	(0.049)	(0.041)	(0.078)	(0.050)
Constant	11.082***	0.089	11.750***	11.306***
	(1.668)	(1.136)	(3.491)	(2.704)
Observations	3,235	3,235	666	666
Test of excluded instruments				
$\chi^2(12)$	78.48	79.77	35.47	26.38
P-value	0.000	0.000	0.000	0.010

Notes: Clustered robust standard errors are reported in parentheses. *** Significant at 1% level, ** Significant at 5% level, * Significant at 10% level. The first two columns correspond to the first- stage estimates of the starting age and schooling equations for the urban sample, and the last two columns correspond to the rural sample.

Table 3.7: IV Ordered Probit Estimates - Second-stage Regression of Health Model (3.2a)

Variables	Urban_1	Urban_2	Urban_3	Rural_1	Rural_2	Rural_3
starting age	0.010	0.009	-0.018	-0.082***	-0.003	0.085***
	(0.011)	(0.008)	(0.019)	(0.021)	(0.007)	(0.018)
years of schooling	-0.032**	-0.028***	0.060***	0.100***	0.004	-0.105***
	(0.013)	(0.006)	(0.018)	(0.033)	(0.009)	(0.029)
age	0.013**	0.011**	-0.024**	0.035	0.001	-0.037*
	(0.006)	(0.005)	(0.011)	(0.022)	(0.003)	(0.022)
age squared	-0.000	-0.000	0.000	-0.000	-0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
male	-0.039**	-0.033*	0.073**	-0.129	-0.002	0.131
	(0.019)	(0.018)	(0.036)	(0.083)	(0.008)	(0.084)
black	-0.013	-0.012	0.026	-0.032	-0.002	0.034
	(0.019)	(0.018)	(0.037)	(0.067)	(0.005)	(0.071)
other race	-0.021	-0.019*	0.040	0.027	0.001	-0.028
	(0.015)	(0.012)	(0.027)	(0.056)	(0.004)	(0.061)
current GDP/capita	-0.009***	-0.008***	0.017***	-0.022*	-0.001	0.023*
	(0.003)	(0.003)	(0.006)	(0.012)	(0.002)	(0.013)
father schooling missing	0.002	0.002	-0.004	0.075	0.003	-0.078
	(0.017)	(0.015)	(0.033)	(0.052)	(0.007)	(0.056)
father illiterate	-0.037	-0.041	0.078	0.229***	-0.013	-0.217***
	(0.028)	(0.034)	(0.062)	(0.072)	(0.013)	(0.061)
father upper primary	0.013	0.011	-0.023	-0.013	-0.001	0.014
	(0.019)	(0.014)	(0.033)	(0.106)	(0.006)	(0.111)
father highschool	0.021	0.017	-0.038	-0.129	-0.015	0.144
	(0.028)	(0.019)	(0.046)	(0.156)	(0.032)	(0.179)
father college	0.016	0.013	-0.029	-0.205	-0.034	0.239
	(0.041)	(0.029)	(0.070)	(0.145)	(0.068)	(0.192)
mother schooling missing	-0.039*	-0.038**	0.077*	0.109	0.005	-0.114
	(0.023)	(0.019)	(0.041)	(0.081)	(0.010)	(0.083)
mother illiterate	0.014	0.012	-0.026	0.133	-0.002	-0.131
	(0.034)	(0.027)	(0.062)	(0.097)	(0.008)	(0.090)
mother upper primary	0.014	0.012	-0.026	0.019	0.001	-0.020
	(0.020)	(0.015)	(0.034)	(0.105)	(0.003)	(0.107)
mother highschool	0.017	0.014	-0.030	0.083	0.000	-0.083
	(0.027)	(0.019)	(0.046)	(0.133)	(0.007)	(0.128)
mother college	0.006	0.005	-0.012	-0.184	-0.029	0.212
	(0.041)	(0.033)	(0.074)	(0.143)	(0.063)	(0.191)
age7_hospital/1000 residents	0.258	0.229	-0.487	3.461	0.147	-3.608
	(0.482)	(0.441)	(0.921)	(2.432)	(0.320)	(2.486)
age7_bed/1000 residents	-0.012*	-0.011	0.023*	-0.002	-0.000	0.002
	(0.007)	(0.007)	(0.014)	(0.033)	(0.001)	(0.034)
age7_doctor/1000 residents	0.065***	0.058***	-0.124***	0.084	0.004	-0.088
	(0.019)	(0.023)	(0.040)	(0.072)	(0.008)	(0.076)
Observations	3,235	3,235	3,235	666	666	666

Notes: Marginal effects are reported rather than probit coefficients. Clustered robust standard errors are reported in parentheses. *** Significant at 1% level, ** Significant at 5% level, * Significant at 10% level. Column 1, 2 and 3 correspond to the marginal effects on the probablity that the health index equals 1, 2 and 3 for the urban residents, and the column 4, 5 and 6 correspond to the marginal effects for the rural residents.

Table 3.8: Probit Estimates of Health Model (3.2b)

Variables	Urban	Rural
starting age	-0.005***	-0.010**
	(0.002)	(0.004)
years of schooling	0.001	0.003
	(0.002)	(0.006)
age	-0.003	0.010
	(0.007)	(0.016)
age squared	0.000	-0.000
	(0.000)	(0.000)
male	-0.039***	-0.016
	(0.014)	(0.036)
black	0.016	-0.024
	(0.029)	(0.050)
other race	0.030*	-0.012
	(0.017)	(0.036)
current GDP/capita	-0.017***	0.002
	(0.004)	(0.010)
father schooling missing	-0.007	0.090**
	(0.017)	(0.039)
father illiterate	0.087*	0.127
	(0.050)	(0.097)
father upper primary	-0.037*	0.141
	(0.023)	(0.141)
father highschool	-0.011	0.026
	(0.030)	(0.118)
father college	-0.037	-0.077
	(0.038)	(0.106)
mother schooling missing	-0.029	0.001
	(0.018)	(0.041)
mother illiterate	0.027	-0.017
	(0.042)	(0.069)
mother upper primary	0.028	-0.064
	(0.028)	(0.084)
mother highschool	-0.010	0.052
	(0.031)	(0.112)
mother college	0.013	0.069
	(0.057)	(0.229)
age7_hospital/1000 residents	-0.790	1.268
	(0.900)	(1.628)
age7_bed/1000 residents	-0.004	-0.045
	(0.012)	(0.028)
age7_doctor/1000 residents	0.061**	0.015
	(0.030)	(0.068)
Observations	3,235	666

Notes: Marginal effects are reported rather than probit coefficients. Clustered robust standard errors are reported in parentheses. *** Significant at 1% level, ** Significant at 5% level, * Significant at 10% level.

Table 3.9: IV Probit Estimates - First-stage Regression of Health Model (3.2b)

Variables	starting age_u	schooling_u	starting age_r	schooling_r
age	0.332***	0.501***	0.188	-0.138
	(0.085)	(0.060)	(0.332)	(0.145)
age squared	-0.005***	-0.006***	-0.002	0.002
	(0.001)	(0.001)	(0.003)	(0.002)
male	-1.985***	-0.876***	-2.427***	-0.995*
	(0.151)	(0.115)	(0.511)	(0.567)
black	-0.617**	-1.126***	-0.873	-0.165
	(0.248)	(0.237)	(1.881)	(1.981)
other race	-0.432***	-1.070***	-0.586*	-0.348
	(0.155)	(0.128)	(0.338)	(0.442)
current GDP/capita	-0.215***	-0.223***	-0.180	-0.002
	(0.046)	(0.040)	(0.445)	(0.126)
father schooling missing	-1.031***	-1.187***	-0.534	-0.932***
	(0.202)	(0.173)	(0.760)	(0.281)
father illiterate	-1.255***	-2.232***	0.518	-1.437
	(0.449)	(0.386)	(1.749)	(0.889)
father upper primary	0.397	0.903***	1.639*	1.767***
	(0.248)	(0.207)	(0.869)	(0.545)
father highschool	1.093***	1.531***	2.472**	3.520***
	(0.272)	(0.252)	(1.100)	(0.996)
father college	1.696***	2.572***	3.436**	4.825***
	(0.340)	(0.277)	(1.457)	(0.864)
mother schooling missing	-0.787***	-2.052***	-1.084*	-1.921***
	(0.179)	(0.157)	(0.601)	(0.324)
mother illiterate	-1.137***	-1.597***	-1.404	-2.268***
	(0.433)	(0.351)	(1.689)	(0.854)
mother upper primary	0.910***	1.044***	1.983*	1.314
	(0.230)	(0.207)	(1.167)	(0.862)
mother highschool	1.502***	1.972***	2.870	1.584
	(0.288)	(0.196)	(4.285)	(2.096)
mother college	1.609***	2.966***	-1.634	0.323
	(0.514)	(0.381)	(2.534)	(2.068)
age7_hospital/1000 residents	9.169	7.687	-0.979	-34.170*
	(8.900)	(6.308)	(55.783)	(18.965)
age7_bed/1000 residents	-0.152	-0.056	-0.407	-0.262
	(0.135)	(0.105)	(0.352)	(0.355)
age7_doctor/1000 residents	1.216***	0.605**	0.512	0.157
	(0.430)	(0.284)	(1.252)	(0.496)
father occupation missing	-0.145	0.106	-0.240	-0.264
	(0.337)	(0.282)	(0.847)	(0.932)
father not work	-0.705***	0.277	0.256	0.088
	(0.218)	(0.204)	(0.742)	(0.843)
father self-employed	-0.390**	0.022	0.141	0.291
	(0.175)	(0.150)	(0.290)	(0.226)
father employer	-0.225	0.825***	0.270	0.762
	(0.329)	(0.254)	(1.681)	(1.329)
mother occupation missing	0.702	0.659	-0.456	0.537
	(0.598)	(0.494)	(0.922)	(1.950)
mother not work	0.721***	0.672***	0.183	0.608
	(0.160)	(0.137)	(0.478)	(0.491)
mother self-employed	0.100	0.530***	0.371	0.878*
	(0.249)	(0.198)	(1.322)	(0.468)
mother employer	-0.351	0.477	-2.356	0.182
	(0.777)	(0.575)	(2.415)	(1.007)
mother unsalaried	-1.597***	-0.415	-1.414	-0.485
	(0.438)	(0.338)	(0.901)	(1.011)
age12_gdp/capita	0.075	0.280***	0.070	0.126
	(0.061)	(0.049)	(0.380)	(0.208)
age7_teacher/school	-0.050	-0.043	0.143*	-0.069
	(0.049)	(0.040)	(0.081)	(0.148)
age11_teacher/school	-0.006	-0.067*	-0.025	0.037
	(0.046)	(0.041)	(0.092)	(0.109)
Constant	11.212***	-0.028	12.743	10.850***
	(1.642)	(1.130)	(8.587)	(3.754)
Observations	3,235	3,235	666	666
Test of excluded instruments				
$\chi^2(12)$	80.32	77.10	91.63	33.47
P-value	0.000	0.000	0.000	0.010

Notes: Clustered robust standard errors are reported in parentheses. *** Significant at 1% level, ** Significant at 5% level, * Significant at 10% level.

Table 3.10: IV Probit Estimates - Second-stage Regression of Health Model (3.2b)

Variables	Urban	Rural
starting age	-0.021***	-0.091***
	(0.008)	(0.018)
years of schooling	-0.014	0.094
	(0.013)	(0.069)
age	0.009	0.036
	(0.009)	(0.023)
age squared	-0.000	-0.000
	(0.000)	(0.000)
male	-0.085***	-0.128
	(0.021)	(0.112)
black	-0.014	-0.061
	(0.029)	(0.065)
other race	0.004	-0.024
	(0.020)	(0.035)
current GDP/capita	-0.022***	-0.018
	(0.004)	(0.037)
father schooling missing	-0.043*	0.072**
	(0.022)	(0.037)
father illiterate	0.020	0.230
	(0.052)	(0.274)
father upper primary	-0.016	0.042
	(0.027)	(0.158)
father highschool	0.035	-0.095
	(0.038)	(0.232)
father college	0.035	-0.151
	(0.055)	(0.331)
mother schooling missing	-0.071**	0.092
	(0.028)	(0.210)
mother illiterate	-0.019	0.084
	(0.042)	(0.132)
mother upper primary	0.059*	0.034
	(0.032)	(0.149)
mother highschool	0.043	0.156
	(0.043)	(0.424)
mother college	0.079	-0.120
	(0.079)	(0.110)
age7_hospital/1000 residents	-0.601	3.212
	(0.961)	(6.224)
age7_bed/1000 residents	-0.006	-0.017
	(0.012)	(0.050)
age7_doctor/1000 residents	0.082**	0.072
	(0.032)	(0.073)
Observations	3,235	666
Overidentification test		
	$\chi^2(10) = 4.427$	$\chi^2(9) = 4.9018$
P-value	0.937	0.843

Notes: Marginal effects are reported rather than probit coefficients. Clustered robust standard errors are reported in parentheses. *** Significant at 1% level, ** Significant at 5% level, * Significant at 10% level. The Amemiya-Lee-Newey statistics for overidentification test of instruments are reported. The overidentification test statistic for the rural sample is distributed as $\chi^2(9)$ since the instrument "mother employer" predicts failure perfectly in the estimation process and is omitted.

Table 3.11a: Quantile Regression Estimates of Schooling Model - Urban Sample

Variables	OLS	q10	q20	q30	q40	q50	q60	q70	q80	q90
starting age	0.227***	0.158***	0.223***	0.258***	0.276***	0.284***	0.254***	0.232***	0.211***	0.162***
	(0.015)	(0.028)	(0.025)	(0.022)	(0.022)	(0.021)	(0.019)	(0.017)	(0.019)	(0.025)
age	0.433***	0.143	0.318***	0.319***	0.389***	0.492***	0.594***	0.631***	0.552***	0.540***
	(0.058)	(0.089)	(0.072)	(0.072)	(0.072)	(0.073)	(0.072)	(0.081)	(0.087)	(0.091)
age squared	-0.005***	-0.002	-0.004***	-0.004***	-0.005***	-0.006***	-0.007***	-0.007***	-0.006***	-0.005***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
male	-0.432***	-0.346**	-0.498***	-0.390***	-0.534***	-0.456***	-0.443***	-0.509***	-0.543***	-0.425**
	(0.118)	(0.176)	(0.150)	(0.143)	(0.136)	(0.145)	(0.144)	(0.153)	(0.179)	(0.177)
black	-1.065***	-0.523*	-0.727**	-0.948***	-0.916***	-1.158***	-1.406***	-1.143***	-1.149***	-1.343***
	(0.234)	(0.313)	(0.331)	(0.309)	(0.246)	(0.266)	(0.303)	(0.378)	(0.339)	(0.378)
other race	-1.054***	-0.774***	-0.818***	-1.041***	-1.000***	-1.037***	-1.036***	-1.074***	-1.123***	-1.277***
	(0.121)	(0.220)	(0.187)	(0.178)	(0.172)	(0.171)	(0.159)	(0.153)	(0.171)	(0.209)
current GDP/capita	-0.189***	-0.002	-0.047	-0.104**	-0.178***	-0.223***	-0.278***	-0.310***	-0.302***	-0.219**
	(0.038)	(0.055)	(0.052)	(0.044)	(0.037)	(0.043)	(0.049)	(0.060)	(0.072)	(0.091)
father schooling missing	-0.952***	-0.807***	-0.728***	-1.030***	-0.901***	-0.989***	-1.191***	-1.263***	-0.914***	-0.625**
	(0.164)	(0.226)	(0.228)	(0.223)	(0.244)	(0.224)	(0.204)	(0.198)	(0.230)	(0.284)
father illiterate	-2.026***	-1.267**	-1.716**	-1.783***	-1.995***	-1.951***	-2.370***	-2.227***	-2.265***	-1.561**
	(0.367)	(0.534)	(0.714)	(0.484)	(0.423)	(0.435)	(0.422)	(0.445)	(0.723)	(0.776)
father upper primary	0.864***	1.384***	1.217***	0.959***	0.860***	0.769***	0.598**	0.488*	0.585**	0.420
	(0.204)	(0.382)	(0.276)	(0.273)	(0.245)	(0.203)	(0.265)	(0.272)	(0.272)	(0.293)
father highschool	1.330***	1.496**	1.574***	1.349***	1.182***	1.100***	0.940***	1.183***	1.114***	0.899***
	(0.243)	(0.738)	(0.465)	(0.293)	(0.245)	(0.252)	(0.276)	(0.337)	(0.340)	(0.319)
father college	2.286***	2.790***	2.543***	2.195***	2.293***	2.001***	1.947***	1.831***	1.920***	1.424***
	(0.259)	(0.549)	(0.427)	(0.479)	(0.405)	(0.383)	(0.314)	(0.356)	(0.320)	(0.332)
mother schooling missing	-1.871***	-1.161***	-1.600***	-2.068***	-2.511***	-2.476***	-2.300***	-1.852***	-1.767***	-1.395***
	(0.158)	(0.226)	(0.216)	(0.201)	(0.236)	(0.246)	(0.207)	(0.210)	(0.200)	(0.278)
mother illiterate	-1.342***	-2.015***	-2.302***	-1.543**	-1.567***	-1.409***	-0.983**	-0.932**	-0.646	-0.696
	(0.360)	(0.537)	(0.788)	(0.661)	(0.498)	(0.529)	(0.490)	(0.386)	(0.492)	(0.588)
mother upper primary	0.800***	1.567***	1.371***	1.069***	0.627***	0.425*	0.390	0.746***	0.761**	0.896***
	(0.198)	(0.532)	(0.358)	(0.273)	(0.237)	(0.226)	(0.284)	(0.257)	(0.318)	(0.301)
mother highschool	1.550***	3.187***	2.422***	1.777***	1.311***	1.188***	1.008***	1.030***	1.287***	1.299***
	(0.194)	(0.499)	(0.363)	(0.271)	(0.238)	(0.229)	(0.203)	(0.292)	(0.310)	(0.353)
mother college	2.318***	3.277***	2.561***	2.387***	2.049***	1.827***	1.963***	2.257***	2.358***	2.024***
	(0.344)	(0.564)	(0.615)	(0.687)	(0.515)	(0.485)	(0.581)	(0.624)	(0.402)	(0.383)
age12_gdp/capita	0.265***	0.082	0.133**	0.166***	0.186***	0.232***	0.325***	0.382***	0.460***	0.404***
	(0.046)	(0.071)	(0.061)	(0.055)	(0.054)	(0.052)	(0.054)	(0.060)	(0.075)	(0.099)
age7_teacher/school	-0.030	0.005	-0.048	-0.015	-0.057	-0.065	-0.038	-0.029	-0.054	-0.005
	(0.038)	(0.080)	(0.058)	(0.052)	(0.053)	(0.052)	(0.043)	(0.045)	(0.055)	(0.067)
age11_teacher/school	-0.066*	-0.072	-0.060	-0.053	-0.028	-0.023	-0.051	-0.107***	-0.084	-0.045
	(0.036)	(0.067)	(0.067)	(0.046)	(0.051)	(0.041)	(0.033)	(0.036)	(0.051)	(0.072)
age7_hospital/1000 residents	6.184	-8.638	-3.857	-3.699	-3.235	7.812	19.009**	21.556***	15.379*	31.903***
	(6.199)	(8.429)	(9.289)	(10.074)	(8.339)	(9.507)	(9.048)	(6.636)	(8.131)	(11.625)
age7_bed/1000 residents	-0.017	0.154	0.029	-0.073	0.035	-0.017	-0.047	-0.049	-0.061	-0.230
	(0.101)	(0.128)	(0.131)	(0.144)	(0.129)	(0.133)	(0.132)	(0.115)	(0.140)	(0.176)
age7_doctor/1000 residents	0.332	0.000	0.321	0.458	0.536	0.580	0.394	0.638*	0.283	0.076
	(0.295)	(0.565)	(0.439)	(0.410)	(0.417)	(0.387)	(0.361)	(0.346)	(0.414)	(0.451)
Constant	-2.177**	-0.504	-2.341	-1.566	-2.211	-3.686***	-4.552***	-4.659***	-2.184	-1.008
	(1.093)	(1.615)	(1.443)	(1.407)	(1.455)	(1.417)	(1.356)	(1.456)	(1.602)	(1.770)
Observations	3,235	3,235	3,235	3,235	3,235	3,235	3,235	3,235	3,235	3,235

Note: Clustered robust standard errors are reported in parentheses. *** Significant at 1% level, ** Significant at 5% level, * Significant at 10% level.

Table 3.11b: Quantile Regression Estimates of Schooling Model - Rural Sample

Variables	sch	q10	q20	q30	q40	q50	q60	q70	q80	q90
starting age	0.163***	0.046	0.094**	0.149***	0.171***	0.178***	0.194***	0.220***	0.203***	0.167**
	(0.034)	(0.033)	(0.039)	(0.042)	(0.045)	(0.051)	(0.044)	(0.043)	(0.057)	(0.072)
age	-0.207	0.056	0.018	-0.248*	-0.225*	-0.181	-0.136	-0.107	-0.271	-0.546*
	(0.131)	(0.144)	(0.159)	(0.138)	(0.119)	(0.153)	(0.187)	(0.202)	(0.231)	(0.300)
age squared	0.003	-0.001	-0.000	0.003*	0.003*	0.002	0.001	0.001	0.004	0.008*
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.004)
male	-0.598**	-0.710***	-0.597**	-0.538**	-0.467*	-0.431	-0.507	-0.485	-0.392	-0.765
	(0.262)	(0.261)	(0.284)	(0.252)	(0.267)	(0.361)	(0.376)	(0.351)	(0.448)	(0.597)
black	-0.090	-0.068	-0.090	0.074	-0.110	-0.120	0.088	-0.478	-0.922	0.041
	(0.411)	(0.421)	(0.443)	(0.438)	(0.463)	(0.468)	(0.440)	(0.464)	(0.585)	(1.169)
other race	-0.246	-0.449	-0.252	-0.223	-0.377	-0.379	-0.012	-0.107	-0.315	0.161
	(0.264)	(0.277)	(0.318)	(0.343)	(0.326)	(0.353)	(0.335)	(0.380)	(0.385)	(0.568)
current GDP/capita	0.076	0.015	0.051	0.115	0.095	0.093	0.149	0.068	-0.005	0.146
	(0.118)	(0.118)	(0.111)	(0.135)	(0.142)	(0.154)	(0.165)	(0.161)	(0.177)	(0.303)
father schooling missing	-0.907***	-1.036***	-1.263***	-1.048***	-1.034***	-0.763**	-0.522	-0.532	-0.662	-0.399
	(0.279)	(0.393)	(0.379)	(0.263)	(0.270)	(0.323)	(0.341)	(0.362)	(0.486)	(0.605)
father illiterate	-1.586***	-1.687***	-2.254***	-1.599**	-1.175**	-1.261**	-1.401**	-1.419*	-1.076	-1.170
	(0.431)	(0.495)	(0.551)	(0.648)	(0.565)	(0.509)	(0.613)	(0.761)	(0.947)	(1.352)
father upper primary	1.325**	1.577	1.456	0.909	1.500	1.751*	1.540*	0.976	0.376	-0.064
	(0.645)	(1.128)	(0.944)	(1.012)	(1.167)	(1.052)	(0.810)	(0.732)	(0.887)	(1.262)
father highschool	3.281***	6.469***	5.365***	4.971***	4.906***	3.479***	2.442***	2.331**	1.730	1.483
	(0.607)	(1.511)	(1.302)	(1.226)	(1.163)	(1.082)	(0.924)	(1.072)	(1.300)	(1.613)
father college	4.171***	5.126***	6.195***	5.641***	5.067***	2.736	4.124***	3.389**	3.384***	3.032*
	(0.906)	(1.912)	(1.745)	(1.754)	(1.908)	(1.763)	(1.536)	(1.391)	(1.261)	(1.584)
mother schooling missing	-1.841***	-0.663	-0.867***	-0.935***	-0.927***	-1.458***	-2.289***	-2.869***	-3.159***	-3.673***
	(0.266)	(0.415)	(0.320)	(0.261)	(0.319)	(0.363)	(0.469)	(0.459)	(0.536)	(0.721)
mother illiterate	-2.087***	-1.272**	-1.612***	-2.029***	-1.420**	-1.841***	-2.403***	-2.606***	-3.041***	-3.763**
	(0.482)	(0.505)	(0.475)	(0.650)	(0.692)	(0.630)	(0.708)	(0.785)	(0.956)	(1.701)
mother upper primary	1.050	-0.633	0.819	0.769	0.670	1.288	1.142	1.592*	0.822	0.688
	(0.647)	(1.305)	(1.193)	(1.186)	(1.236)	(1.264)	(1.171)	(0.931)	(0.823)	(1.533)
mother highschool	0.961	1.291	0.451	0.528	0.802	2.009	1.767	0.901	0.200	-0.520
	(0.717)	(1.541)	(1.479)	(1.593)	(1.703)	(1.414)	(1.107)	(0.810)	(0.704)	(1.124)
mother college	0.514	0.308	-0.695	0.759	0.553	1.949	1.949	0.675	2.921	1.688
	(1.660)	(2.127)	(2.125)	(2.417)	(2.672)	(2.706)	(2.690)	(2.955)	(3.073)	(3.662)
age12_gdp/capita	0.049	0.078	-0.026	-0.049	-0.067	-0.019	-0.029	0.098	0.200	0.110
	(0.125)	(0.147)	(0.136)	(0.147)	(0.158)	(0.188)	(0.199)	(0.183)	(0.183)	(0.291)
age7_teacher/school	-0.055	-0.054	-0.041	-0.047	-0.090	-0.061	-0.043	-0.090	-0.095	-0.048
	(0.062)	(0.084)	(0.067)	(0.061)	(0.071)	(0.100)	(0.112)	(0.106)	(0.132)	(0.151)
age11_teacher/school	0.024	0.019	0.001	0.036	-0.000	0.010	0.049	0.036	0.085	0.079
	(0.066)	(0.053)	(0.061)	(0.068)	(0.066)	(0.093)	(0.111)	(0.117)	(0.128)	(0.145)
age7_hospital/1000 residents	-40.534***	-8.684	-6.555	-26.266	-31.472*	-35.516*	-42.604**	-29.642	-36.080*	-49.910*
	(14.243)	(14.654)	(15.826)	(16.192)	(17.694)	(20.156)	(21.206)	(20.082)	(19.314)	(27.315)
age7_bed/1000 residents	-0.181	-0.162	-0.215	-0.030	-0.010	-0.168	-0.263	-0.305	-0.276	-0.089
	(0.229)	(0.233)	(0.238)	(0.226)	(0.247)	(0.307)	(0.312)	(0.272)	(0.298)	(0.457)
age7_doctor/1000 residents	0.018	0.772	1.001	0.142	0.445	0.117	0.115	0.138	-0.401	-1.668
	(0.579)	(0.670)	(0.677)	(0.609)	(0.614)	(0.660)	(0.704)	(0.647)	(0.822)	(1.312)
Constant	10.290***	1.650	3.088	7.778***	8.555***	9.142***	9.304**	9.318**	13.672***	20.274***
	(2.525)	(2.905)	(3.101)	(2.917)	(2.664)	(3.321)	(3.768)	(3.918)	(4.542)	(5.789)
Observations	666	666	666	666	666	666	666	666	666	666

Note: Clustered robust standard errors are reported in parentheses. *** Significant at 1% level, ** Significant at 5% level, * Significant at 10% level.

Table 3.12: IV Ordered Probit Estimates - Second-stage Regression of Health Model (3.2a) with Full Sample

Variables	Urban_1	Urban_2	Urban_3	Rural_1	Rural_2	Rural_3
starting age	-0.005	-0.003	0.009	-0.032	-0.007***	0.039*
	(0.007)	(0.004)	(0.011)	(0.020)	(0.003)	(0.023)
years of schooling	-0.028***	-0.018***	0.046***	0.017	0.004	-0.021
	(0.010)	(0.005)	(0.015)	(0.051)	(0.010)	(0.061)
age	0.013***	0.009***	-0.022***	-0.002	-0.000	0.002
	(0.005)	(0.003)	(0.007)	(0.009)	(0.002)	(0.011)
age squared	-0.000	-0.000*	0.000	0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
male	-0.068***	-0.043***	0.111***	-0.144***	-0.025***	0.169***
	(0.012)	(0.009)	(0.020)	(0.028)	(0.009)	(0.028)
black	-0.011	-0.007	0.018	0.043	0.006*	-0.049
	(0.015)	(0.010)	(0.025)	(0.047)	(0.003)	(0.049)
other race	0.000	0.000	-0.001	0.049*	0.011**	-0.060**
	(0.012)	(0.008)	(0.020)	(0.025)	(0.005)	(0.029)
current GDP/capita	-0.008***	-0.005***	0.014***	-0.025***	-0.006***	0.031***
	(0.002)	(0.001)	(0.003)	(0.008)	(0.002)	(0.009)
father schooling missing	-0.014	-0.009	0.023	0.040	0.009	-0.048
	(0.015)	(0.009)	(0.024)	(0.049)	(0.009)	(0.058)
father illiterate	-0.021	-0.015	0.035	0.054	0.007***	-0.061
	(0.022)	(0.017)	(0.039)	(0.086)	(0.002)	(0.088)
father upper primary	0.003	0.002	-0.005	-0.006	-0.001	0.008
	(0.014)	(0.008)	(0.022)	(0.080)	(0.020)	(0.100)
father highschool	0.020	0.011	-0.031	0.021	0.004	-0.025
	(0.024)	(0.012)	(0.036)	(0.144)	(0.020)	(0.164)
father college	0.007	0.004	-0.012	-0.108	-0.061	0.170
	(0.033)	(0.019)	(0.052)	(0.137)	(0.127)	(0.264)
mother schooling missing	-0.032*	-0.021**	0.053*	0.034	0.008	-0.042
	(0.017)	(0.011)	(0.028)	(0.075)	(0.015)	(0.091)
mother illiterate	-0.018	-0.013	0.031	-0.000	-0.000	0.000
	(0.024)	(0.018)	(0.041)	(0.075)	(0.017)	(0.092)
mother upper primary	0.030*	0.017**	-0.047*	-0.028	-0.008	0.037
	(0.018)	(0.008)	(0.026)	(0.055)	(0.020)	(0.074)
mother highschool	0.046*	0.023***	-0.070**	0.058	0.006**	-0.064
	(0.026)	(0.009)	(0.034)	(0.112)	(0.003)	(0.112)
mother college	0.021	0.012	-0.033	-0.116	-0.070	0.186
	(0.036)	(0.017)	(0.054)	(0.132)	(0.132)	(0.265)
age7_hospital/1000 residents	0.374	0.238	-0.612	-1.575	-0.354	1.930
	(0.331)	(0.215)	(0.545)	(1.308)	(0.372)	(1.655)
age7_bed/1000 residents	-0.016***	-0.010***	0.026***	0.019	0.004	-0.023
	(0.005)	(0.004)	(0.009)	(0.016)	(0.003)	(0.020)
age7_doctor/1000 residents	0.067***	0.043***	-0.110***	-0.006	-0.001	0.008
	(0.014)	(0.011)	(0.024)	(0.034)	(0.008)	(0.042)
Observations	6,439	6,439	6,439	1,573	1,573	1,573

Notes: Marginal effects are reported rather than probit coefficients. Clustered robust standard errors are reported in parentheses. *** Significant at 1% level, ** Significant at 5% level, * Significant at 10% level.

Table 3.13: IV Probit estimates - Second-stage Regression of Health Model (3.2b) with Full Sample

Variables	Urban	Rural
starting age	-0.019***	-0.057***
	(0.006)	(0.006)
years of schooling	-0.019*	0.035***
	(0.011)	(0.013)
age	0.015**	0.011
	(0.006)	(0.011)
age squared	-0.000*	-0.000
	(0.000)	(0.000)
male	-0.087***	-0.122***
	(0.014)	(0.026)
black	-0.021	0.012
	(0.023)	(0.041)
other race	0.017	-0.000
	(0.016)	(0.025)
current GDP/capita	-0.018***	-0.013
	(0.003)	(0.008)
father schooling missing	-0.032*	0.046
	(0.018)	(0.028)
father illiterate	-0.021	0.081
	(0.034)	(0.056)
father upper primary	0.020	0.038
	(0.020)	(0.093)
father highschool	0.033	-0.040
	(0.032)	(0.089)
father college	0.051	-0.028
	(0.047)	(0.129)
mother schooling missing	-0.061***	0.024
	(0.020)	(0.034)
mother illiterate	-0.015	0.023
	(0.031)	(0.049)
mother upper primary	0.063**	-0.005
	(0.025)	(0.070)
mother highschool	0.096***	0.121
	(0.037)	(0.090)
mother college	0.155**	-0.085
	(0.061)	(0.110)
age7_hospital/1000 residents	0.502	0.731
	(0.559)	(1.167)
age7_bed/1000 residents	-0.015*	-0.021
	(0.009)	(0.017)
age7_doctor/1000 residents	0.075***	-0.014
	(0.024)	(0.045)
Observations	6,439	1,573
Overidentification test		
	$\chi^2(10) = 7.143$	$\chi^2(10) = 13.690$
P-value	0.712	0.188

Notes: Marginal effects are reported rather than probit coefficients. Clustered robust standard errors are reported in parentheses. *** Significant at 1% level, ** Significant at 5% level, * Significant at 10% level. The Amemiya-Lee-Newey statistics for overidentification test of instruments are reported.

Table 3.14a: Quantile Regression Estimates of Schooling Model -Urban Full Sample

Variables	OLS	q10	q20	q30	q40	q50	q60	q70	q80	q90
starting age	0.222***	0.142***	0.187***	0.235***	0.257***	0.280***	0.256***	0.242***	0.222***	0.219***
	(0.011)	(0.013)	(0.017)	(0.015)	(0.017)	(0.015)	(0.016)	(0.014)	(0.016)	(0.018)
age	0.385***	0.141***	0.194***	0.303***	0.367***	0.430***	0.460***	0.528***	0.486***	0.505***
	(0.038)	(0.048)	(0.054)	(0.046)	(0.055)	(0.054)	(0.056)	(0.055)	(0.064)	(0.074)
age squared	-0.005***	-0.002***	-0.003***	-0.004***	-0.005***	-0.006***	-0.006***	-0.007***	-0.006***	-0.005***
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
male	0.040	-0.115	0.011	0.043	0.024	0.063	0.068	0.101	0.001	0.021
	(0.085)	(0.091)	(0.094)	(0.091)	(0.101)	(0.106)	(0.116)	(0.111)	(0.129)	(0.145)
black	-0.946***	-0.268	-0.628***	-0.982***	-0.830***	-0.820***	-1.019***	-0.928***	-0.902***	-1.190***
	(0.170)	(0.229)	(0.213)	(0.249)	(0.245)	(0.219)	(0.217)	(0.284)	(0.233)	(0.335)
other race	-0.968***	-0.709***	-0.888***	-0.977***	-0.935***	-0.930***	-0.899***	-0.902***	-0.961***	-0.999***
	(0.091)	(0.112)	(0.113)	(0.105)	(0.125)	(0.121)	(0.122)	(0.112)	(0.122)	(0.153)
current GDP/capita	-0.107***	0.014	-0.028	-0.065**	-0.100***	-0.117***	-0.134***	-0.151***	-0.149***	-0.142*
	(0.024)	(0.028)	(0.027)	(0.029)	(0.027)	(0.028)	(0.027)	(0.032)	(0.040)	(0.073)
father schooling missing	-1.122***	-0.934***	-0.754***	-0.995***	-1.117***	-1.306***	-1.423***	-1.329***	-1.296***	-1.064***
	(0.104)	(0.125)	(0.124)	(0.118)	(0.142)	(0.150)	(0.162)	(0.155)	(0.163)	(0.165)
father illiterate	-1.488***	-1.183***	-1.385***	-1.182***	-1.310***	-1.840***	-1.772***	-1.805***	-1.548***	-1.006*
	(0.262)	(0.199)	(0.407)	(0.331)	(0.274)	(0.346)	(0.358)	(0.343)	(0.565)	(0.560)
father upper primary	0.778***	0.869***	1.060***	0.985***	0.748***	0.658***	0.623***	0.618***	0.637***	0.754***
	(0.145)	(0.240)	(0.211)	(0.207)	(0.205)	(0.183)	(0.174)	(0.199)	(0.183)	(0.253)
father highschool	1.561***	1.851***	1.958***	1.922***	1.574***	1.214***	1.327***	1.392***	1.287***	1.043***
	(0.177)	(0.376)	(0.349)	(0.277)	(0.225)	(0.226)	(0.197)	(0.228)	(0.260)	(0.243)
father college	2.483***	2.708***	2.694***	2.836***	2.683***	2.423***	2.477***	2.364***	1.963***	1.573***
	(0.202)	(0.335)	(0.371)	(0.319)	(0.345)	(0.284)	(0.268)	(0.248)	(0.228)	(0.263)
mother schooling missing	-1.665***	-1.062***	-1.051***	-1.373***	-1.834***	-2.104***	-2.223***	-2.158***	-1.727***	-1.548***
	(0.110)	(0.121)	(0.121)	(0.115)	(0.151)	(0.160)	(0.157)	(0.154)	(0.169)	(0.186)
mother illiterate	-1.461***	-1.950***	-1.481***	-1.513***	-1.648***	-1.686***	-1.424***	-1.310***	-0.891**	-0.725**
	(0.253)	(0.276)	(0.449)	(0.270)	(0.398)	(0.383)	(0.345)	(0.350)	(0.405)	(0.348)
mother upper primary	0.917***	1.036***	1.466***	1.168***	1.006***	0.814***	0.553***	0.519**	0.736***	0.876***
	(0.158)	(0.254)	(0.220)	(0.198)	(0.194)	(0.180)	(0.205)	(0.205)	(0.236)	(0.263)
mother highschool	1.823***	3.156***	2.956***	2.286***	1.894***	1.580***	1.288***	1.230***	1.322***	1.315***
	(0.145)	(0.297)	(0.300)	(0.214)	(0.196)	(0.182)	(0.171)	(0.214)	(0.210)	(0.265)
mother college	2.465***	3.292***	3.054***	2.754***	2.317***	2.308***	2.139***	1.995***	2.092***	1.989***
	(0.276)	(0.350)	(0.451)	(0.406)	(0.455)	(0.392)	(0.349)	(0.326)	(0.307)	(0.362)
age12_gdp/capita	0.191***	0.091**	0.075	0.138***	0.180***	0.185***	0.211***	0.218***	0.292***	0.313***
	(0.033)	(0.044)	(0.046)	(0.046)	(0.044)	(0.043)	(0.039)	(0.042)	(0.049)	(0.069)
age7_teacher/school	-0.011	0.054	0.028	-0.010	-0.013	-0.025	-0.034	-0.010	-0.041	-0.015
	(0.031)	(0.046)	(0.042)	(0.042)	(0.043)	(0.043)	(0.038)	(0.038)	(0.042)	(0.042)
age11_teacher/school	-0.060*	-0.076*	-0.050	-0.063	-0.059	-0.040	-0.053*	-0.053*	-0.075*	-0.037
	(0.031)	(0.045)	(0.048)	(0.049)	(0.043)	(0.040)	(0.030)	(0.028)	(0.039)	(0.046)
age7_hospital/1000 residents	1.899	-0.782	-8.921*	-4.737	-3.043	3.691	2.471	7.158	8.674	14.318*
	(4.229)	(4.550)	(5.044)	(5.364)	(6.334)	(5.391)	(5.732)	(6.639)	(6.147)	(7.811)
age7_bed/1000 residents	0.055	0.153**	0.153*	0.132*	0.059	0.014	0.061	0.025	0.039	-0.085
	(0.067)	(0.074)	(0.080)	(0.073)	(0.095)	(0.086)	(0.083)	(0.095)	(0.107)	(0.119)
age7_doctor/1000 residents	0.003	-0.441	-0.271	-0.102	0.133	0.116	0.079	0.063	0.087	-0.100
	(0.205)	(0.289)	(0.281)	(0.281)	(0.295)	(0.270)	(0.235)	(0.268)	(0.289)	(0.377)
Constant	-1.643**	-0.619	-0.559	-1.875**	-2.473**	-2.916***	-2.268**	-2.707**	-1.353	-0.954
	(0.750)	(0.936)	(1.056)	(0.906)	(1.077)	(1.034)	(1.074)	(1.068)	(1.197)	(1.430)
Observations	6,451	6,451	6,451	6,451	6,451	6,451	6,451	6,451	6,451	6,451

Note: Clustered robust standard errors are reported in parentheses. *** Significant at 1% level, ** Significant at 5% level,
* Significant at 10% level.

Table 3.14b: Quantile Regression Estimates of Schooling Model-Rural Full Sample

Variables	OLS	q10	q20	q30	q40	q50	q60	q70	q80	q90
starting age	0.153***	0.021	0.053**	0.097***	0.121***	0.127***	0.150***	0.195***	0.220***	0.244***
	(0.020)	(0.022)	(0.025)	(0.022)	(0.020)	(0.023)	(0.036)	(0.034)	(0.036)	(0.044)
age	0.004	0.024	0.039	-0.107	-0.056	-0.004	0.023	0.050	0.022	-0.015
	(0.074)	(0.065)	(0.072)	(0.082)	(0.066)	(0.057)	(0.081)	(0.094)	(0.108)	(0.152)
age squared	-0.000	-0.000	-0.001	0.001	0.000	-0.000	-0.001	-0.001	-0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
male	-0.088	-0.183	-0.302**	-0.338**	-0.039	0.037	0.029	-0.011	-0.077	-0.111
	(0.130)	(0.148)	(0.147)	(0.150)	(0.131)	(0.130)	(0.178)	(0.191)	(0.206)	(0.310)
black	-0.493*	-0.530**	-0.624**	-0.616	-0.456*	-0.442	-0.165	-0.311	-0.489	-0.537
	(0.268)	(0.250)	(0.278)	(0.376)	(0.253)	(0.277)	(0.352)	(0.329)	(0.397)	(0.712)
other race	-0.321*	-0.396**	-0.296	-0.298*	-0.283*	-0.346**	-0.334*	-0.144	-0.060	-0.064
	(0.168)	(0.181)	(0.181)	(0.176)	(0.170)	(0.145)	(0.200)	(0.206)	(0.253)	(0.351)
current GDP/capita	0.020	0.115*	0.186***	0.190**	0.139*	0.102	0.083	0.035	-0.007	-0.140
	(0.071)	(0.065)	(0.068)	(0.086)	(0.072)	(0.068)	(0.088)	(0.080)	(0.079)	(0.139)
father schooling missing	-1.054***	-0.661***	-0.889***	-0.966***	-0.941***	-0.848***	-0.903***	-0.785***	-1.280***	-1.113**
	(0.155)	(0.192)	(0.190)	(0.166)	(0.135)	(0.125)	(0.184)	(0.211)	(0.317)	(0.458)
father illiterate	-1.424***	-0.794***	-1.158***	-1.599***	-1.286***	-1.066***	-1.247***	-1.259***	-1.074*	-1.130*
	(0.275)	(0.249)	(0.250)	(0.346)	(0.332)	(0.302)	(0.387)	(0.439)	(0.575)	(0.627)
father upper primary	1.404**	0.700	0.665	1.129	1.453*	1.539*	1.347	1.402	1.114	2.419**
	(0.563)	(0.559)	(0.783)	(0.753)	(0.763)	(0.818)	(0.877)	(0.956)	(1.048)	(1.017)
father highschool	2.789***	4.468***	4.176***	3.353***	3.021***	3.054***	2.404**	2.516***	1.960*	2.798*
	(0.591)	(1.541)	(0.954)	(1.095)	(1.092)	(0.993)	(0.991)	(0.799)	(1.024)	(1.458)
father college	4.042***	7.607***	5.128***	4.139**	4.122**	4.175**	4.296***	3.855***	3.268**	1.974
	(0.932)	(2.249)	(1.830)	(1.653)	(1.647)	(1.622)	(1.649)	(1.449)	(1.498)	(1.878)
mother schooling missing	-1.617***	-1.313***	-1.211***	-0.992***	-0.887***	-0.944***	-1.122***	-2.040***	-2.408***	-2.570***
	(0.158)	(0.203)	(0.168)	(0.173)	(0.146)	(0.130)	(0.228)	(0.312)	(0.387)	(0.494)
mother illiterate	-1.498***	-1.468***	-1.533***	-1.246***	-1.055***	-0.819**	-0.799**	-1.566***	-1.885***	-2.331***
	(0.283)	(0.243)	(0.297)	(0.431)	(0.397)	(0.335)	(0.386)	(0.451)	(0.526)	(0.726)
mother upper primary	0.906**	0.639	0.364	0.806	0.751	1.821**	1.999**	1.310	1.395**	0.629
	(0.456)	(0.675)	(0.454)	(0.559)	(0.820)	(0.900)	(0.878)	(0.931)	(0.634)	(0.813)
mother highschool	1.905***	0.960	2.552*	2.668**	2.909**	2.806**	2.955***	1.628*	0.939	0.180
	(0.614)	(1.824)	(1.457)	(1.307)	(1.248)	(1.164)	(1.047)	(0.844)	(0.721)	(0.913)
mother college	3.309*	-0.408	2.789	2.571	4.751	4.212	4.919*	5.004**	4.017*	4.869*
	(1.813)	(2.936)	(3.121)	(3.212)	(2.994)	(2.738)	(2.725)	(2.466)	(2.402)	(2.487)
age12_gdp/capita	0.105	0.040	0.035	-0.052	-0.019	-0.001	0.044	0.132	0.203*	0.280*
	(0.078)	(0.070)	(0.071)	(0.086)	(0.079)	(0.079)	(0.097)	(0.105)	(0.106)	(0.156)
age7_teacher/school	0.001	0.036	0.034	-0.003	-0.031	0.009	0.019	0.056	-0.039	-0.129
	(0.058)	(0.049)	(0.047)	(0.053)	(0.055)	(0.066)	(0.084)	(0.103)	(0.129)	(0.160)
age11_teacher/school	0.041	-0.007	-0.028	0.002	0.011	-0.017	-0.012	-0.017	0.104	0.236
	(0.047)	(0.030)	(0.040)	(0.045)	(0.034)	(0.039)	(0.056)	(0.091)	(0.117)	(0.146)
age7_hospital/1000 residents	-7.311	0.762	9.146	1.441	-1.636	0.309	-1.991	-7.254	-1.123	-13.619
	(8.860)	(7.975)	(7.947)	(9.584)	(8.982)	(8.409)	(11.198)	(11.392)	(11.467)	(15.675)
age7_bed/1000 residents	-0.202	-0.097	-0.255*	-0.085	-0.096	-0.203*	-0.227	-0.317**	-0.334**	-0.163
	(0.128)	(0.150)	(0.136)	(0.141)	(0.123)	(0.111)	(0.142)	(0.142)	(0.159)	(0.229)
age7_doctor/1000 residents	-0.275	0.188	0.381	-0.100	0.088	0.223	0.002	-0.174	-0.407	-1.200*
	(0.310)	(0.375)	(0.401)	(0.328)	(0.316)	(0.303)	(0.344)	(0.426)	(0.455)	(0.633)
Constant	4.830***	1.511	1.613	4.946***	4.234***	3.934***	4.158**	4.793**	5.958***	8.335***
	(1.471)	(1.338)	(1.444)	(1.690)	(1.413)	(1.198)	(1.723)	(1.891)	(2.236)	(2.980)
Observations	1,578	1,578	1,578	1,578	1,578	1,578	1,578	1,578	1,578	1,578

Note: Clustered robust standard errors are reported in parentheses. *** Significant at 1% level, ** Significant at 5% level,
* Significant at 10% level.

Figure 1.1: Distributions of the Big Five Personality Traits (%)

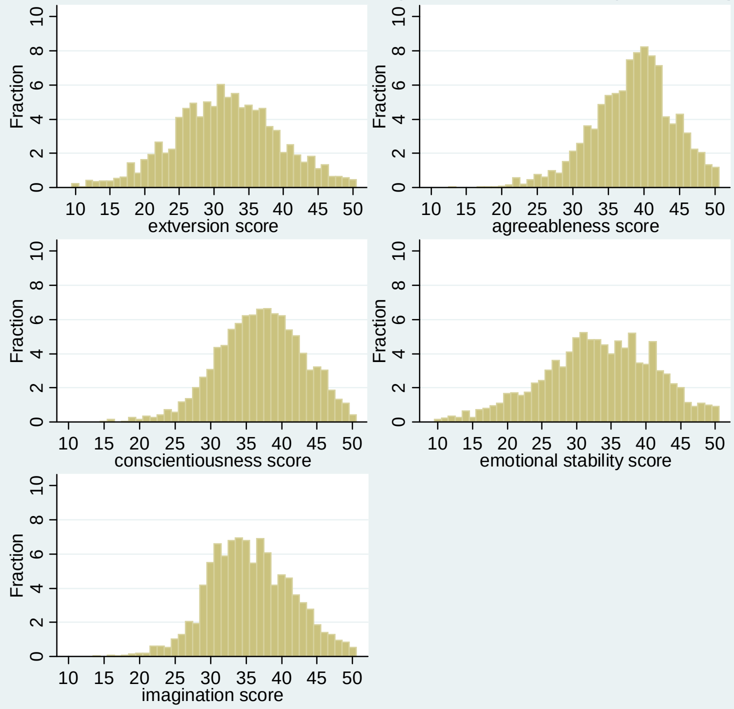


Figure 1.2: Distributions of Teacher's Ratings for Child Behavior (%)

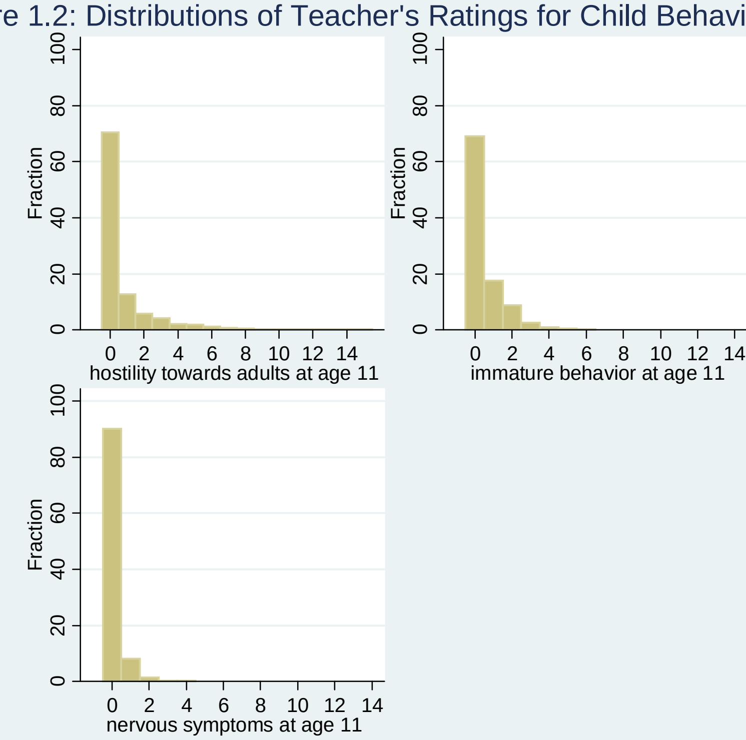



Figure 3.1: Distribution of Age Started to Work (%)
0
2
4
6
8
10
12
14
16
5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31
Age started to work
Urban
Rural

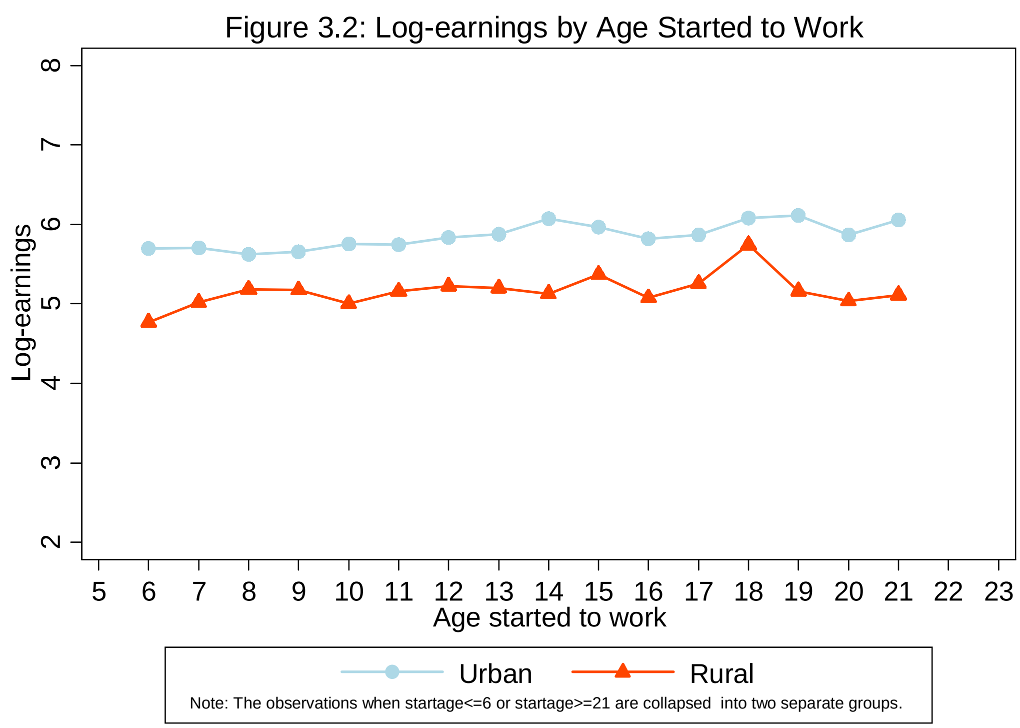


Figure 3.2: Log-earnings by Age Started to Work
Log-earnings
2
3
4
5
6
7
8
5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23
Age started to work
Urban
Rural
Note: The observations when startage<=6 or startage>=21 are collapsed into two separate groups.

Figure 3.3: Health Score by Age Started to Work

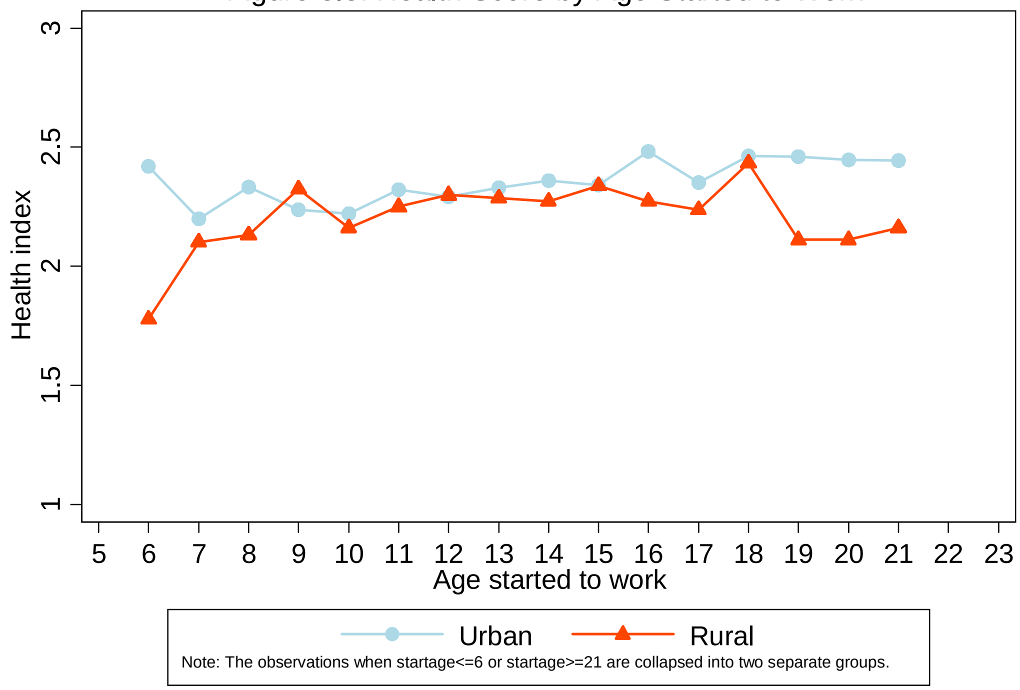


Figure 3.4a: Years of Schooling by Age Started to Work

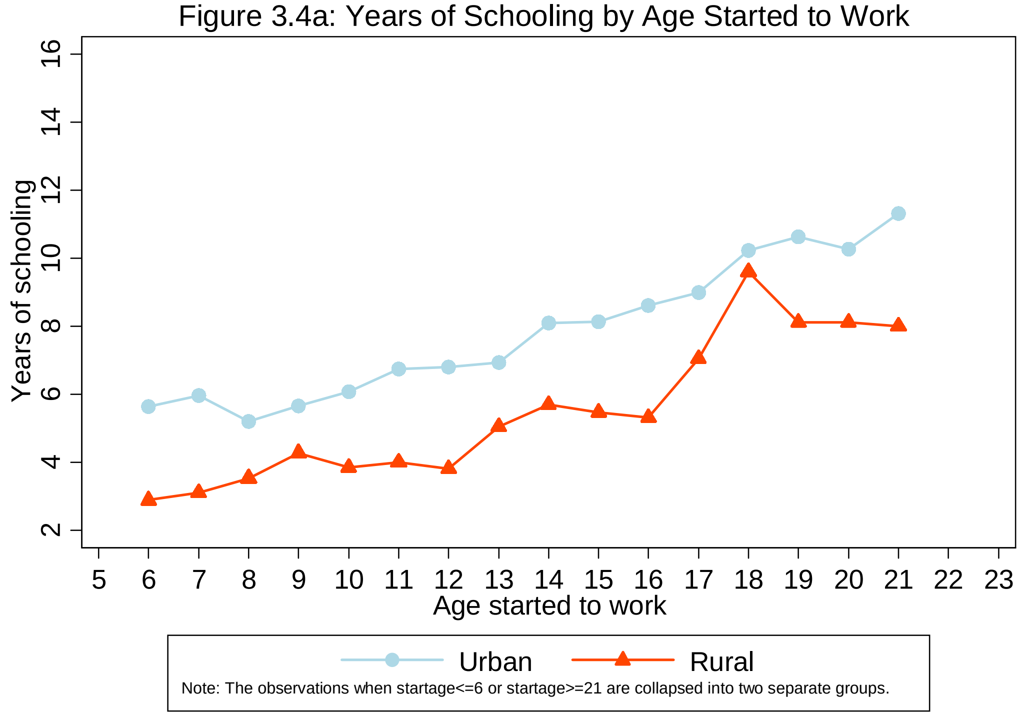
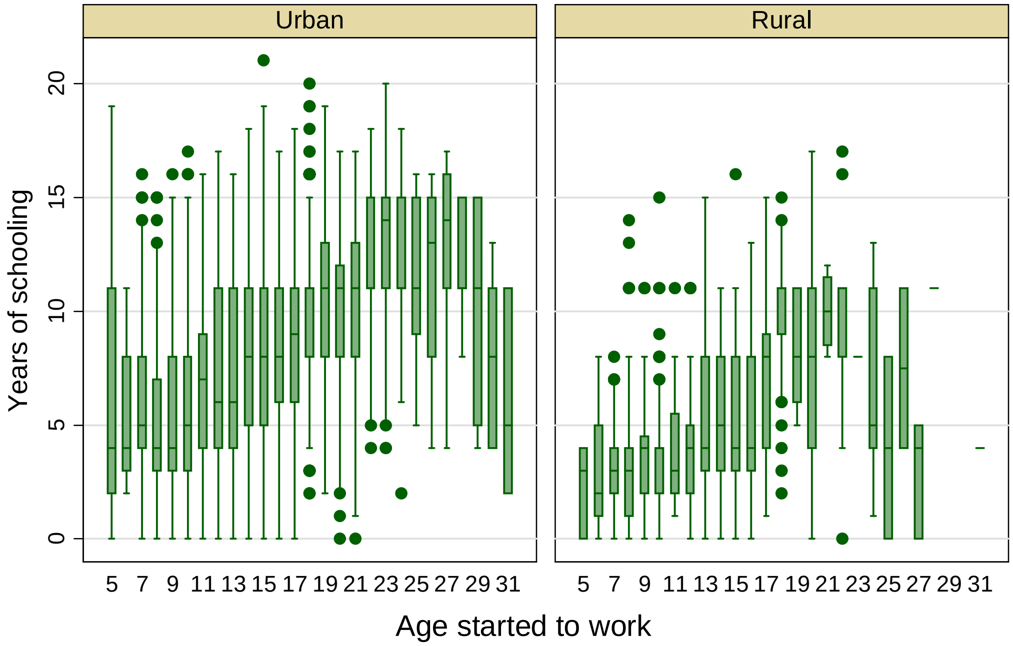


Figure 3.4b: Years of Schooling by Age Started to Work



Graphs by areas

Appendix

Table A.1.1 Descriptions of Occupations

Occupation Categories (Abbreviations in bold)	Socioeconomic group	Examples
Managerial occupation category (occupation=1)	Managers in central and local government, industry, commerce, etc. - large establishments Employers in industry, commerce, etc. -small establishments Managers in industry, commerce, etc. -small establishments Farmers: employers & managers	Marketing and sales manager research manager, office manager Hotel and accommodation manager, transport and distribution manager Leisure and sports manager, customer care manager Farm manager
Non-manual and technical occupation category (occupation=2)	Intermediate non-manual: Ancillary workers and artists Junior non-manual workers Intermediate non-manual: Foremen and supervisors Farmers: own account	Artist, author, arts officer Draughts person, photographer Fire service officer, database assistant Farmer
Professional occupation category (occupation=3)	Professional workers: Self-employed Professional workers: Employees	Mechanical engineer, judge, lawyer Psychologist, actuary
Manual occupation category (occupation=4)	Foremen & supervisors: manual Skilled manual workers Own account workers: non-professionals Personal service workers Semi-skilled manual workers Agricultural workers Unskilled manual workers	Van driver, paramedic Pipe fitter, printer, baker, butcher Barber, plasterer Kitchen and catering assistants, bar staff Rail construction and maintenance Farm worker, forestry worker Road sweeper

Table A.1.2 The Big Five Personality Traits

Extraversion	Agreeableness
I don't talk a lot. I keep in the background. I have little to say. I don't like to draw attention to myself. I am quiet around strangers. I am the life of the party. I feel comfortable around people. I start conversations. I talk to a lot of different people at parties. I don't mind being the center of attention.	I feel little concern for others. I insult people. I am not interested in other people's problems. I am not really interested in others. I am interested in people. I sympathize with others' feelings. I have a soft heart. I take time out for others. I feel others' emotions. I make people feel at ease.
Conscientiousness	Emotional Stability
I leave my belongings around. I make a mess of things. I often forget to put things back in their proper place. I shirk my duties. I am always prepared. I pay attention to details. I get chores done right away. I like order. I follow a schedule. I am exacting in my work.	I get stressed out easily. I worry about things. I am easily disturbed. I get upset easily. I change my mood a lot. I have frequent mood swings. I get irritated easily. I often feel blue. I am relaxed most of the time. I seldom feel blue.
Imagination	
I have difficulty understanding abstract ideas. I am not interested in abstract ideas. I do not have a good imagination. I have a rich vocabulary. I have a vivid imagination. I have excellent ideas. I am quick to understand things. I use difficult words. I spend time reflecting on things. I am full of ideas.	

Table A.1.3 Correlations among the Big Five

	extraversion	agreeableness	conscientiousness	emotional stability	imagination
extraversion	1.0000				
agreeableness	0.3651	1.0000			
conscientiousness	0.1323	0.2406	1.0000		
emotional stability	0.2750	0.1518	0.2293	1.0000	
imagination	0.3842	0.3371	0.2263	0.1244	1.0000

Table A.3.1: Ordered Probit Estimates of Health Model (2a) with Full Sample

Variables	Urban_1	Urban_2	Urban_3	Rural_1	Rural_2	Rural_3
starting age	-0.002***	-0.002***	0.004***	-0.006***	-0.002***	0.008***
	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.003)
years of schooling	-0.008***	-0.006***	0.014***	-0.009***	-0.002***	0.012***
	(0.001)	(0.001)	(0.002)	(0.003)	(0.001)	(0.004)
age	0.005	0.004	-0.009	-0.005	-0.001	0.007
	(0.003)	(0.002)	(0.006)	(0.008)	(0.002)	(0.011)
age squared	0.000	0.000	-0.000	0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
male	-0.056***	-0.040***	0.096***	-0.103***	-0.023***	0.125***
	(0.007)	(0.005)	(0.011)	(0.017)	(0.005)	(0.020)
black	0.011	0.007	-0.019	0.042	0.007**	-0.050
	(0.013)	(0.008)	(0.022)	(0.038)	(0.003)	(0.041)
other race	0.021***	0.015***	-0.036***	0.045**	0.012**	-0.056**
	(0.008)	(0.005)	(0.013)	(0.020)	(0.006)	(0.025)
current GDP/capita	-0.006***	-0.004***	0.011***	-0.020***	-0.005***	0.025***
	(0.002)	(0.001)	(0.003)	(0.006)	(0.002)	(0.008)
father schooling missing	0.016*	0.011*	-0.027*	0.031	0.008	-0.040
	(0.009)	(0.006)	(0.015)	(0.019)	(0.005)	(0.024)
father illiterate	0.018	0.011	-0.029	0.012	0.003	-0.015
	(0.020)	(0.012)	(0.031)	(0.035)	(0.007)	(0.043)
father upper primary	-0.015	-0.012	0.027	0.003	0.001	-0.004
	(0.011)	(0.009)	(0.021)	(0.052)	(0.013)	(0.064)
father highschool	-0.022	-0.017	0.039	0.035	0.006	-0.041
	(0.014)	(0.013)	(0.027)	(0.077)	(0.007)	(0.084)
father college	-0.050***	-0.047**	0.097***	-0.088	-0.051	0.138
	(0.015)	(0.019)	(0.034)	(0.079)	(0.078)	(0.157)
mother schooling missing	0.008	0.006	-0.014	0.007	0.002	-0.009
	(0.009)	(0.006)	(0.015)	(0.021)	(0.005)	(0.026)
mother illiterate	0.018	0.012	-0.030	-0.014	-0.004	0.018
	(0.022)	(0.013)	(0.035)	(0.038)	(0.013)	(0.051)
mother upper primary	0.006	0.004	-0.011	-0.040	-0.015	0.056
	(0.013)	(0.008)	(0.021)	(0.046)	(0.024)	(0.069)
mother highschool	-0.003	-0.002	0.005	0.016	0.004	-0.020
	(0.015)	(0.011)	(0.025)	(0.073)	(0.013)	(0.087)
mother college	-0.036	-0.031	0.067	-0.062	-0.029	0.090
	(0.022)	(0.024)	(0.046)	(0.109)	(0.077)	(0.186)
age7_hospital/1000 residents	0.392	0.279	-0.670	-1.421	-0.374	1.796
	(0.312)	(0.223)	(0.535)	(1.040)	(0.289)	(1.323)
age7_bed/1000 residents	-0.018***	-0.013***	0.031***	0.011	0.003	-0.014
	(0.005)	(0.003)	(0.008)	(0.015)	(0.004)	(0.019)
age7_doctor/1000 residents	0.065***	0.046***	-0.112***	-0.010	-0.003	0.013
	(0.012)	(0.009)	(0.020)	(0.033)	(0.009)	(0.042)
Observations	6,439	6,439	6,439	1,573	1,573	1,573

Notes: Marginal effects rather than ordered probit estimates are reported. Clustered robust standard errors are reported in parentheses. *** Significant at 1% level, ** Significant at 5% level, * Significant at 10% level.

Table A.3.2: IV Ordered Probit Estimates - First-stage Regression of Health Model (2a) with Full Sample

Variables	starting age_u	schooling_u	starting age_r	schooling_r
age	0.247***	0.434***	0.164	0.031
	(0.060)	(0.040)	(0.102)	(0.074)
age squared	-0.004***	-0.006***	-0.002*	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)
male	-1.775***	-0.351***	-2.136***	-0.493***
	(0.113)	(0.082)	(0.220)	(0.128)
black	-0.487**	-0.953***	-0.760*	-0.538**
	(0.203)	(0.171)	(0.397)	(0.275)
other race	-0.253**	-0.931***	-0.306	-0.352**
	(0.118)	(0.093)	(0.219)	(0.177)
current GDP/capita	-0.167***	-0.131***	-0.259***	-0.014
	(0.032)	(0.025)	(0.090)	(0.070)
father schooling missing	-0.903***	-1.296***	-0.700***	-1.077***
	(0.146)	(0.107)	(0.234)	(0.160)
father illiterate	-1.133***	-1.700***	0.302	-1.324***
	(0.319)	(0.277)	(0.360)	(0.284)
father upper primary	0.439**	0.868***	1.315**	1.683***
	(0.187)	(0.145)	(0.660)	(0.562)
father highschool	1.316***	1.827***	2.459***	3.023***
	(0.194)	(0.182)	(0.890)	(0.586)
father college	1.750***	2.753***	3.101**	4.422***
	(0.271)	(0.215)	(1.440)	(0.967)
mother schooling missing	-0.984***	-1.871***	-0.617***	-1.669***
	(0.129)	(0.110)	(0.228)	(0.161)
mother illiterate	-0.917***	-1.660***	-0.977***	-1.642***
	(0.290)	(0.254)	(0.361)	(0.285)
mother upper primary	0.648***	1.073***	1.210	0.997**
	(0.182)	(0.164)	(0.749)	(0.490)
mother highschool	1.570***	2.208***	3.733***	2.702***
	(0.204)	(0.153)	(0.890)	(0.636)
mother college	1.782***	3.163***	0.067	3.590*
	(0.377)	(0.309)	(1.669)	(1.896)
age7_hospital/1000 residents	3.295	2.758	-15.228	-8.149
	(6.290)	(4.077)	(11.207)	(9.305)
age7_bed/1000 residents	-0.087	0.035	0.009	-0.218*
	(0.101)	(0.071)	(0.189)	(0.132)
age7_doctor/1000 residents	0.769**	0.193	-0.110	-0.307
	(0.352)	(0.214)	(0.465)	(0.351)
father occupation missing	0.072	0.215	0.046	-0.100
	(0.270)	(0.202)	(0.760)	(0.483)
father not work	-0.444***	0.125	-0.180	-0.196
	(0.169)	(0.137)	(0.361)	(0.256)
father self-employed	-0.717***	-0.040	-0.556**	0.106
	(0.139)	(0.103)	(0.257)	(0.152)
father employer	-0.416*	0.890***	0.232	0.841*
	(0.250)	(0.226)	(0.733)	(0.451)
mother occupation missing	0.383	0.827**	-0.335	1.470**
	(0.532)	(0.388)	(0.884)	(0.591)
mother not work	0.857***	0.743***	0.437	0.687**
	(0.109)	(0.101)	(0.349)	(0.331)
mother self-employed	-0.007	0.505***	0.188	0.842***
	(0.167)	(0.129)	(0.400)	(0.299)
mother employer	-0.067	1.135***	-0.185	3.604*
	(0.495)	(0.439)	(1.021)	(2.042)
mother unsalaried	-1.536***	-0.344	-1.453***	-0.173
	(0.259)	(0.220)	(0.401)	(0.317)
age12_gdp/capita	-0.007	0.187***	0.200**	0.147**
	(0.049)	(0.035)	(0.097)	(0.074)
age7_teacher/school	-0.022	-0.025	0.163**	0.014
	(0.037)	(0.033)	(0.077)	(0.061)
age11_teacher/school	0.018	-0.053	-0.147***	0.028
	(0.038)	(0.034)	(0.055)	(0.054)
Constant	12.492***	0.659	12.790***	6.146***
	(1.195)	(0.759)	(2.090)	(1.533)
Observations	6,439	6,439	1,573	1,573

Notes: Clustered robust standard errors are reported in parentheses. *** Significant at 1% level, ** Significant at 5% level, * Significant at 10% level.

Table A.3.3: Probit Estimates of Health Model(2b) with Full Sample

Variables	Urban	Rural
starting age	-0.005***	-0.005*
	(0.001)	(0.003)
years of schooling	-0.001	-0.008**
	(0.002)	(0.004)
age	0.005	0.006
	(0.005)	(0.011)
age squared	-0.000	-0.000
	(0.000)	(0.000)
male	-0.059***	-0.034
	(0.010)	(0.021)
black	0.003	0.015
	(0.022)	(0.042)
other race	0.039***	-0.009
	(0.012)	(0.025)
current GDP/capita	-0.015***	-0.003
	(0.003)	(0.008)
father schooling missing	0.004	0.047*
	(0.013)	(0.026)
father illiterate	0.028	0.013
	(0.034)	(0.044)
father upper primary	-0.004	0.045
	(0.018)	(0.100)
father highschool	-0.023	-0.053
	(0.024)	(0.081)
father college	-0.031	-0.020
	(0.031)	(0.127)
mother schooling missing	-0.016	-0.014
	(0.013)	(0.027)
mother illiterate	0.029	0.015
	(0.031)	(0.046)
mother upper primary	0.035*	-0.051
	(0.021)	(0.065)
mother highschool	0.033	0.003
	(0.026)	(0.092)
mother college	0.074	0.035
	(0.046)	(0.206)
age7_hospital/1000 residents	0.453	1.507
	(0.547)	(1.224)
age7_bed/1000 residents	-0.017**	-0.042**
	(0.008)	(0.018)
age7_doctor/1000 residents	0.066***	-0.015
	(0.023)	(0.046)
Observations	6,439	1,573

Notes: Marginal effects are reported rather than probit coefficients. Clustered robust standard errors are reported in parentheses. *** Significant at 1% level, ** Significant at 5% level, * Significant at 10% level.

Table A.3.4: IV Probit Estimates - first-stage Regression of Health Model (2b) with Full Sample

Variables	starting age_u	sching_u	sarting age_r	schooling_r
age	0.250***	0.439***	0.146	0.036
	(0.060)	(0.040)	(0.099)	(0.073)
age squared	-0.004***	-0.006***	-0.002	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)
male	-1.775***	-0.350***	-2.137***	-0.494***
	(0.112)	(0.082)	(0.216)	(0.125)
black	-0.485**	-0.953***	-0.737*	-0.543**
	(0.200)	(0.170)	(0.381)	(0.263)
other race	-0.253**	-0.931***	-0.291	-0.355**
	(0.117)	(0.093)	(0.211)	(0.172)
current GDP/capita	-0.171***	-0.135***	-0.233***	-0.020
	(0.032)	(0.025)	(0.085)	(0.069)
father schooling missing	-0.898***	-1.289***	-0.691***	-1.078***
	(0.145)	(0.107)	(0.227)	(0.154)
father illiterate	-1.132***	-1.698***	0.314	-1.329***
	(0.320)	(0.279)	(0.344)	(0.279)
father upper primary	0.440**	0.868***	1.273**	1.696***
	(0.187)	(0.145)	(0.625)	(0.590)
father highschool	1.319***	1.828***	2.459***	3.025***
	(0.193)	(0.183)	(0.869)	(0.561)
father college	1.752***	2.756***	3.013**	4.452***
	(0.267)	(0.214)	(1.375)	(0.917)
mother schooling missing	-0.985***	-1.869***	-0.627***	-1.666***
	(0.128)	(0.110)	(0.221)	(0.160)
mother illiterate	-0.913***	-1.656***	-0.988***	-1.639***
	(0.286)	(0.254)	(0.342)	(0.275)
mother upper primary	0.646***	1.068***	1.187	1.002**
	(0.181)	(0.164)	(0.768)	(0.488)
mother highschool	1.562***	2.192***	3.702***	2.718***
	(0.199)	(0.153)	(0.889)	(0.616)
mother college	1.767***	3.140***	0.009	3.616**
	(0.379)	(0.309)	(1.438)	(1.740)
age7_hospital/1000 residents	3.537	3.071	-11.209	-9.498
	(6.208)	(4.069)	(10.933)	(9.015)
age7_bed/1000 residents	-0.090	0.030	-0.074	-0.193
	(0.101)	(0.071)	(0.182)	(0.130)
age7_doctor/1000 residents	0.766**	0.195	-0.426	-0.207
	(0.341)	(0.214)	(0.455)	(0.336)
father occupation missing	0.055	0.191	-0.040	-0.084
	(0.259)	(0.209)	(0.647)	(0.499)
father not work	-0.496***	0.057	-0.024	-0.229
	(0.162)	(0.140)	(0.362)	(0.246)
father self-employed	-0.710***	-0.049	-0.356	0.053
	(0.135)	(0.103)	(0.235)	(0.155)
father employer	-0.399	0.927***	0.409	0.800*
	(0.248)	(0.225)	(0.670)	(0.439)
mother occupation missing	0.433	0.782**	-0.014	1.350**
	(0.502)	(0.399)	(0.707)	(0.587)
mother not work	0.834***	0.705***	0.236	0.766***
	(0.105)	(0.102)	(0.266)	(0.188)
mother self-employed	-0.035	0.481***	0.003	0.897***
	(0.163)	(0.134)	(0.355)	(0.284)
mother employer	-0.152	1.042**	-0.554	3.764***
	(0.471)	(0.454)	(0.819)	(1.456)
mother unsalaried	-1.577***	-0.401*	-1.652***	-0.099
	(0.256)	(0.220)	(0.298)	(0.212)
age12_gdp/capita	0.000	0.196***	0.146*	0.161**
	(0.049)	(0.035)	(0.089)	(0.077)
age7_teacher/school	-0.021	-0.022	0.235***	-0.006
	(0.036)	(0.031)	(0.070)	(0.062)
age11_teacher/school	0.016	-0.058*	-0.115**	0.015
	(0.036)	(0.033)	(0.054)	(0.046)
Constant	12.453***	0.604	13.049***	6.064***
	(1.191)	(0.762)	(2.001)	(1.508)
Observations	6,439	6,439	1,573	1,573
Test of excluded instruments				
$\chi^2(12)$	206.06	120.52	142.27	56.36
P-value	0.000	0.000	0.000	0.000

Notes: Clustered robust standard errors are reported in parentheses. *** Significant at 1% level, ** Significant at 5% level, * Significant at 10% level.

References

Almlund, M., A. L. Duckworth, J. J. Heckman, and T. Kautz (2011). Personality Psychology and Economics. *In Eric A. Hanushek, Stephen Machin and Ludger Woessmann (Eds.), Handbook of the Economics of Education, Vol. 4 (pp. 1-181).* Amsterdam: Elsevier.

Assaad, R., D. Levison, and N. Zibani (2001). The Effect of Child Work on School Enrollment in Egypt. *Economic Research Forum, Working Papers No.111.*

Barrick, M.R., M.K. Mount, and R. Gupta (2003). Meta-Analysis of the Relationship between the Five-Factor-Model of Personality and Holland's Occupational Types. *Personnel Psychology 56*, 45-74.

Bedi, A. S. and J. H. Y. Edwards (2002). The Impact of School Quality on Earnings and Educational Returns–Evidence from a Low-income Country. *Journal of Development Economics 68*, 157-185.

Beegle, K., R. Dehejia, and R. Gatti (2005). Why should We Care about Child Labor? The Education, Labor Market, and Health Consequences of Child Labor. *World Bank Policy Research Working Paper 3479.*

Berecki-Gisolf, J., J. Lucke, R. Hockey, and A. Dobson (2008). Transitions into Informal Caregiving and Out of Paid Employment of Women in Their 50s. *Social Science and Medicine 67*, 122-127.

Bertrand, M. and S. Mullainathan (2001). Do People Mean What They Say? Implications for Subjective Survey Data. *American Economic Review 91*, 67-72.

Bittman, M., T. Hill and C. Thomson (2007). The Impact of Caring on Informal Carers' Employment, Income and Earnings: a Longitudinal Approach. *Australian Journal of Social Issues 42*, 255-272.

Boaz, R. F. and C. F. Muller (1992). Paid Work and Unpaid Help by Caregivers of the Disabled and Frail Elders. *Medical Care 30(2),* 149-158.

Bolin, K., B. Lindgren, and P. Lundborg (2008). Your Next of Kin or Your Own Career? Caring and Working among the 50+ of Europe. *Journal of Health Economics 27,* 718-738.

Boozer, M. A. and T. K. Suri (2001). Child Labor and Schooling Decisions in Ghana. *Working Paper, Yale University.*

Bouchard, T. J. and J. C. Loehlin (2001). Genes, evolution, and personality. *Behavior Genetics 31(3)*, 243-273.

Bound, J., D. A. Jaeger and R. Baker (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.

Butler, J. S., R. V. Burkhauser, J. M. Mitchell, and T. P. Pincus (1987). Measurement Error in Self-Reported Health Variables. *The Review of Economics and Statistics 4*, 644-650.

Campbell, R. J. (2004). Campbell's Psychiatric Dictionary. New York: Oxford University Press.

Carneiro, P., C. Crawford and A. Goodman (2007). The Impact of Early Cognitive and Non-Cognitive Skills on Later Outcomes. *CEE DP 92,* Centre for the Economics of Education, London School of Economics, London, UK.

Caspi, A. (2000). The Child is Father of the Man: Personality Continuities from Childhood to Adulthood. *Journal of Personality and Social Psychology 78(1),* 158-172.

Costa, P. T. and R. R. McCrae (1992a). Four Ways Five Factors are Basic. *Personality and Individual Differences 13(6)*, 653-665.

Costa, P. T. and R. R. McCrae (1992b). NEO Personality Inventory Professional Manual. Odessa, FL: Psychological Assessment Resources.

Costa, P. T. and R. R. McCrae (1994). Set Like Plaster? Evidence for the Stability of Adult Personality. *In Todd F. Heatherton and Joel Lee Weinberger (Eds.), Can Personality Change? (pp. 21-40).* Washington, DC: American Psychological Association.

Couch, K. A., M. C. Daly and D. A. Wolf (1999). Time? Money? Both? The Allocation of Resources to Older Parents. *Demography 36*, 219-232.

Do, Y. K. (2008). The Effect of Informal Caregiving on Labor Market Outcomes in South Korea. *Asia Health Policy Program Working Paper #1, Shorenstein Asia-Pacific Research Center, Stanford University.*

Doty, P., M. E. Jackson, and W. Crown (1998). The Impact of Female Caregivers' Employment Status on Patterns of Formal and Informal Eldercare. *The Gerontologist 38(3)*, 331-341.

Drago, F. (2011). Self-esteem and Earnings. *Journal of Economic Psychology 32*, 480-488.

Duncan, G. J. and R. Dunifon (1998). Soft-skills and Long-run Labor Market Success. *Research in Labor Economics 17,* 123-149.

Emerson, P. M. and A. P. Souza (2007). Is Child Labor Harmful? The Impact of Working Earlier in Life on Adult Earnings. *IZA Discussion Paper No. 3027*.

Engquist, G., C. Johnson, A. Lind and L.P. Barnette (2010). Medicaid-funded Long-term Care: Toward More Home- and Community-Based Options. Center for Health care Strategies, Inc.

Epstein, C. F., C. Seron, B. Oglensky, and R. Saute (1999). The Part-Time Paradox: Time Norms, Professional Life, Family and Gender. New York: Routledge.

Ettner, S. L. (1995). The Impact of "Parent Care" on Female Labor Supply Decisions. *Demography 32(1),* 63-80.

Ettner, S. L. (1996). The Opportunity Costs of Elder Care. *Journal of Human Resources 31(1),* 189-205.

Fassa, A. G., L. A. Facchini, M. M. D. Agnol, and D. Christiani (2000). Child Labor and Health: Problems and Perspectives. *International Journal of Occupational and Environmental Health 6(1)*, 55-62.

Ferreira, E. S. (2001). Trabalho Infantil: Historia Situacao Atual. Canoas: ULBRA.

Fletcher, J. M. (2012). The Effects of Personality Traits on Adult Labor Market Outcomes: Evidence from Siblings. *IZA Discussion Paper No. 6391.*

Fronstin, P., D. H. Greenberg, and P. K. Robins (2005). The Labor Market Consequences of Childhood Maladjustment. *Social Science Quarterly, Supplement to Vol. 86,* 1170-1195.

Georgetown University Law Center (2010). Flexible Work Arrangements: A Definition and Examples. Workplace Flexibility 2010, Georgetown University Law Center.

Ghodsian, M. (1977). Children's Behaviour and the BSAG: Some Theoretical and Statistical Considerations. *British Journal of Social and Clinical Psychology 16(1),* 23-28.

Goldberg, L. R. (1993). The Structure of Phenotypic Personality Traits. *American Psychologist 48(1),* 26-34.

Goldberg, L. R. (1999). A Broad-bandwidth, Public Domain, Personality Inventory Measuring the Lower Level Facets of Several Five-factor Models. *In I. Mervielde, I. Deary, F. De Fruyt and F. Ostendorf (Eds.), Personality Psychology in Europe, Vol. 7 (pp. 7-28).* Tilburg, The Netherlands: Tilburg University Press.

Gottschalk, P. (2005). Can Work Alter Welfare Recipients' Reliefs? *Journal of Policy Analysis and Management 24(3),* 485-498.

Graitcer, P. and L. Lerer (2000). The Impact of Child Labor on Health. Report of a Field Investigation in Egypt, Washington DC, World Bank.

Groves, M. O. (2005). How Important is Your Personality? Labor Market Returns to Personality for Women in the US and UK. *Journal of Economic Psychology 26,* 827-841.

Gunnarsson, V., P. F. Orazem, and M. A. Sánchez (2006). Child Labor and School Achievement in Latin America. *The World Bank Economic Review 20(1)*, 31-54.

Halpern, D. F. (2005). How Time-flexible Work Policies can Reduce Stress, Improve Health and Save Money. *Stress and Health 21,* 157-168.

Ham, R., P. N. R. Junankar, and R. Wells (2009). Antagonistic Managers, Careless Workers and Extraverted Salespeople: An Examination of Personality in Occupational Choice. *IZA Discussion Paper 4193.*

Heady, C. (2003). The Effect of Child Labor on Learning Achievement. *World Development 31(2),* 385?98.

Heckman, J. J., L. Malofeeva, R. Pinto, and P. A. Savelyev (2010). Understanding the Mechanisms through Which an Influential Early Childhood Program Boosted Adult Outcomes. *Unpublished manuscript.* Department of Economics, University of Chicago.

Heckman, J. J., S. H. Moon, R. Pinto, P. Savelyev, and A. Yavitz (2010). Analyzing Social Experiments as Implemented: A Reexamination of the Evidence from the Highscope Perry Preschool Program. *Quantitative Economics 1(1)*, 1-46.

Heckman, J. J., J. Stixrud, and S. Urzua (2006). The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior. J*ournal of Labor Economics 24(3)*, 411-482.

Heineck, G. and S. Anger (2010). The Returns to Cognitive Abilities and Personality Traits in Germany. *Labour Economics 17,* 535-546.

Heitmueller, A. (2007). The Chicken or the Egg? Endogeneity in Labour Market Participation of Informal Carers in England. *Journal of Health Economics 26,* 536-559.

Heitmueller, A. and K. Inglis (2007). The Earnings of Informal Carers: Wage Differentials and Opportunity Costs. *Journal of Health Economics 26*, 821-841.

Helson, R., C. Jones, and V. S. Y. Kwan (2002). Personality Change over 40 years of Adulthood: Hierarchical Linear Modeling Analyses of Two Longitudinal Samples. *Journal of Personality and Social Psychology 83(3),* 752-766.

Heywood, J. S., W. S. Sieberty, and X. Weiz (2007). The Implicit Wage Costs of Family Friendly Work Practices. *Oxford Economic Papers 59,* 275-300.

Hill, E. J., A. J. Hawkins, M. Ferris, and M. Weitzman (2001). Finding an Extra Day a Week: The Positive Influence of Perceived Job Flexibility on Work and Family Life Balance. *Family Relations 50(1)*, 49-58.

IBGE (2007). Child labor 2001. Instituto Brasileiro de Geografia e Estatistica.

IBGE, DPE, and DEPIS (1997). Interviewer manual. Living Standards Measurement Study Survey. Instituto Brasileiro de Geografia e Estatistica, Diretoria De Pesqiisas and Departamento De Populacao E Indicadores Sociais.

ICFTU (2004). Internationally Recognised Core Labour Standards in Brazil. Report for the WTO General Council Review of the Trade Policies of Brazil, International Confederation of Free Trade Unions, 2004.

Ilahi, N., P. F. Orazem, and G. Sedlacek (2001). The Implications of Child Labor for Adult Wages, Income and Poverty: Retrospective Evidence from Brazil. *World Bank, unpublished working paper.*

International Labour Organisation (2010). Accelerating Action against Child Labour. Report of the Director-General, International Labour Conference, 99th session.

Jackson, M. (2006). Personality Traits and Occupational Attainment. European Sociological Review 22(2), 187-199.

John, O. P. and S. Srivastava (1999). The Big Five Trait Taxonomy: History, Measurement and Theoretical Perspectives. *In L. A. Pervin and Oliver P. John (Eds.), Handbook of Personality: Theory and Research (2nd ed., pp. 102-139).* New York: Guilford Press.

Johnson, J. A. (1983). Criminality, Creativity, and Craziness: Structural Similarities in Three Types of Nonconformity. *In William Slaufer and James M Day (Eds.), Personality Theory, Moral Developments and Criminal Behavior (pp. 81-105).* New York: Lexington.

Judge, T. A., C. A. Higgins, C. J. Thoresen, and M. R. Barrick (1999). The Big Five Personality Traits, General Mental Ability, and Career Success across the Life Span. *Personnel Psychology 52*, 621-652.

Kalpan, G. A. and T. Camacho (1983). Perceived Health and Mortality: A Nine Year Follow-up on the Human Population Laboratory Cohort. *American Journal of Epidemiology 117*, 292-298.

Kassouf, A. L., M. Mckee, and E. Mossialos (2001). Early Entrance to the Job Market and Its Effect on Adult Health: Evidence from Brazil. *Health Policy and Planning 16(1)*, 21-28.

Kossek, E. E. and C. Ozeki (1999). Bridging the Work-Family Policy and Productivity Gap: A Literature Review. *Community, Work and Family 2(1)*, 7-32.

Latif, E. (2006). Labour supply effects of informal caregiving in Canada. *Canadian Public Policy 32:* 413-429.

Lee, C. and P. F. Orazem (2010). Lifetime Health Consequences of Child Labor in Brazil. *In Randall K.Q. Akee, Eric V. Edmonds, Konstantinos Tatsiramos (ed.) Child Labor and the Transition between School and Work (Research in Labor Economics, Volume 31, pp. 99-133).* Emerald Group Publishing Limited.

Liu, L., X. Dong and X. Zheng (2010). Parental Care and Married Women's Labor Supply in Urban China. *Feminist Economics 16(3)*, 169-192.

McCrae, R. R. and P. T. Costa (1994). The Stability of Personality: Observation and Evaluations. *Current Directions in Psychological Science 3(6),* 173-175.

McCrate, E. (2005). Flexible Hours, Workplace Authority, and Compensating Wage Differentials in the US. *Feminist Economics 11(1)*, 11-39.

McGee, D. L., Y. Liao, G. Cao, and R. S. Cooper (1999). Self-reported Health Status and Mortality in a Multiethnic US Cohort. *American Journal of Epidemiology 149(1),* 41-46.

Michaud, P. C., A. Heitmueller, and Z. Nazarov (2010). A Dynamic Analysis of Informal Care and Employment in England. *Labour Economics 17,* 455-465.

Miilunpalo, S., I. Vuori, P. Oja, M. Pasanen, and H. Urponen (1997). Self-rated Health Status as a Health Measure: The Predictive Value of Self-reported Health Status on the Use of Physician Services and on Mortality in the Working-age Population. *Journal of Clinical Epidemiology 50(5)*, 517-528.

Milcent, C., J. Huguenin, and D. C. Machado (2005). Children in Brazil: Health, Education and Work. *Anais do XXXIII Encontro Nacional de Economia [Proceedings of the 33th Brazilian Economics Meeting], ANPEC, No. 173.*

Moura, Esmeralda Blanco B. de. (1982). Mulheres e menores no trabalho industrial: os fatores sexo e idade na dinamica do capital. Petropolis: Vozes.

Moutafi, J., A. Furnham, and J. Crump (2003). Demographic and Personality Predictors of Intelligence: A Study Using the Neo Personality Inventory and the Myers-briggs Type Indicator. *European Journal of Personality 17,* 79-94.

Moutafi, J., A. Furnham, and L. Paltiel (2004). Why is Conscientiousness Negatively Correlated with Intelligence? *Personality and Individual Differences 37,* 1013-1022.

Mueller, G. and E. J. Plug (2006). Estimating the Effect of Personality on Male and Female Earnings. *Industrial and Labor Relations Review 60(1),* 3-22.

Nandi, A. and C. Nicoletti (2009). Explaining Personality Pay Gaps in the UK. ISER working paper No. 2009-22, provided by Institute for Social and Economic Research.

Nolte, M. A. and M. A. Servais (2010). Occupation and Industry Coding in HRS/AHEAD. Health and Retirement Study Documentation Report DR-021.

Nyhus, E. K. and E. Pons (2005). The Effects of Personality on Earnings. *Journal of Economic Psychology 26*, 363-384.

O'Donnell, O., F. C. Rosati and E. Doorslaer (2005). Health Effects of Child Work: Evidence from Rural Vietnam. *Journal of Population Economics 18*: 437-467.

Ozer, D. J. and S. P. Reise (1994). Personality Assessment. *Annual Review of Psychology 45*, 357-388.

Parikh, A. and E. Sadoulet (2005). The Effect of Parents' Occupation on Child Labor and School Attendance in Brazil. *UC Berkeley: Department of Agricultural and Resource Economics, UCB. CUDARE Working Paper No. 1000.*

Pavalko, E. K. and K. A. Henderson (2006). Combining Care Work and Paid Work Do Workplace Policies Make a Difference? *Research on Aging 28(3),* 359-374.

Powdthavee, N., C. J. Boyce, and A. M. Wood (2011). Can Money Change Who We Are? Estimating the Effects of Unearned Income on Measures of Incentive-enhancing Personality Traits. IZA Discussion Paper No. 6131.

Psacharopoulos, G. (1997). Child Labor Versus Educational Attainment: Some Evidence from Latin America. *Journal of Population Economics 10(4),* 377-386.

Ray, R. and G. Lancaster (2005). The Impact of Children's Work on Schooling: Multi-Country Evidence Based on SIMPOC Data. *International Labour Review 144(2)*, 189-210.

Rhoads, S. E. (1993). Incomparable Worth: Pay Equity Meets the Market. Cambridge: Cambridge University Press.

Roberts, B. W., K. E. Walton, and W. Viechtbauer (2006). Patterns of Mean-level Change in Personality Traits across the Life Course: A Meta-analysis of Longitudinal Studies. *Psychological Bulletin 132(1),* 1-25.

Rotter, J. B. (1966). Generalized Expectancies for Internal Versus External Control of Reinforcement. *Psychological Monographs: General and Applied 80(1),* 1-28.

Semykina, A. and S. J. Linz (2007). Gender Differences in Personality and Earnings: Evidence from Russia. *Journal of Economic Psychology 28,* 387-410.

Srivastava, S., O. P. John, S. D. Gosling, and J. Potter (2003). Development of Personality in Early and Middle Adulthood: Set like Plaster or Persistent Change? *Journal of Personality and Social Psychology 84(5),* 1041-1053.

Sutin, A. R., P. T. C. Jr., R. Miech, and W. W. Eaton (2009). Personality and career success: Concurrent and longitudinal relations. European Journal of Personality 23, 71-84.

University of London, Institute of Education, Center for Longitudinal Studies (2008). *National Child Development Study: Childhood Data, Sweeps 0-3, 1958-1974 [Computer File]. 2nd Edition.* Colchester, Essex: UK Data Archive [distributor]. National Birthday Trust Fund, National Children's Bureau, [original data producer(s)]. SN: 5565.

University of London, Institute of Education, Center for Longitudinal Studies (2010). *National Child Development Study: Sweep 8, 2008-2009: First Deposit [Computer File]. 2nd Edition.* Colchester, Essex: UK Data Archive [distributor]. SN: 6137.

U.S. Department of Health and Human Services (1997). Active Aging: A Shift in the Paradigm. Washington, DC: U.S. Department of Health and Human Services, Assistant Secretary for Planning and Evaluation, Office of Disability, Aging and Long-Term Care Policy.

U.S. Department of Health and Human Services (2011). A Profile of Older Americans: 2011. Washington, DC: U.S. Department of Health and Human Services, Administration on Aging.

Van Houtven, C. H., N. B. Coe, and M. Skira (2010). Effect of Informal Care on Work, Wages and Wealth. *Working Papers, Center for Retirement Research at Boston College wp2010-22, Center for Retirement Research, revised Dec 2010.*

Wakabayashi, C. and K. M. Donato (2006). Does Caregiving Increase Poverty among Women in Later Life? Evidence from the Health and Retirement Survey. *Journal of Health and Social Behavior 47(3),* 258-274.

Vita

Xuelian Zhang

2013	Ph.D in Economics, Rutgers University, New Brunswick, New Jersey
2011	M.A in Economics, Rutgers University, New Brunswick, New Jersey