2010

ESSAYS ON HUMAN CAPITAL, HEALTH CAPITAL, AND THE LABOR MARKET

Charles Hokayem
University of Kentucky, charles.hokayem@uky.edu

ABSTRACT OF DISSERTATION

CHARLES HOKAYEM

The Graduate School

University of Kentucky

2010

ESSAYS ON HUMAN CAPITAL, HEALTH CAPITAL, AND THE LABOR MARKET

ABSTRACT OF DISSERTATION

A dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy in the
College of Business and Economics
at the University of Kentucky

By

Charles Hokayem

Lexington, Kentucky

Director: Dr. James P. Ziliak, Carol Martin Gatton Chair in Microeconomics and

Director of the University of Kentucky Center for Poverty Research

University of Kentucky

Lexington, Kentucky

2010

ABSTRACT OF DISSERTATION

ESSAYS ON HUMAN CAPITAL, HEALTH CAPITAL, AND THE LABOR MARKET

This dissertation consists of three essays concerning the effects of human capital and health capital on the labor market. Chapter 1 presents a structural model that incorporates a health capital stock to the traditional learning-by-doing model. The model allows health to affect future wages by interrupting current labor supply and on-the-job human capital accumulation. Using data on sick time from the Panel Study Income of Dynamics the model is estimated using a nonlinear Generalized Method of Moments estimator. The results show human capital production exhibits diminishing returns. Health capital production increases with the current stock of health capital, or better current health improves future health. Among prime age working men, the effect of health on human capital accumulation is relatively small. Chapter 2 explores the role of another form of human capital, noncognitive skills, in explaining racial gaps in wages. Chapter 2 adds two noncognitive skills, locus of control and self-esteem, to a simple wage specification to determine the effect of these skills on the racial wage gap (white, black, and Hispanic) and the return to these skills across the wage distribution. The wage specifications are estimated using pooled, between, and quantile estimators. Results using the National Longitudinal Survey of Youth 1979 show these skills account for differing portions of the racial wage gap depending on race and gender. Chapter 3 synthesizes the idea of health and on-the-job human capital accumulation from Chapter 1 with the idea of noncognitive skills in Chapter 2 to examine the influence of these skills on human capital and health capital accumulation in adult life. Chapter 3 introduces noncognitive skills to a life cycle labor supply model with endogenous health and human capital accumulation. Noncognitive skills, measured by degree of future orientation, self-efficacy, trust-hostility, and aspirations, exogenously affect human capital and health production. The model uses noncognitive skills assessed in the early years of the Panel Study of Income Dynamics and relates these skills to health and human capital accumulation during adult life. The main findings suggest individuals with high self-efficacy receive higher future wages.

KEYWORDS: Health Capital, Human Capital, Labor Supply, Noncognitive Skills, Generalized Method of Moments

Charles Hokayem

Date

ESSAYS ON HUMAN CAPITAL, HEALTH CAPITAL, AND THE LABOR MARKET

By

Charles Hokayem

_Dr. James P. Ziliak________
Director of Dissertation

Dr. Christopher R. Bollinger
Director of Graduate Studies

____

RULES FOR THE USE OF DISSERTATIONS

Unpublished dissertations submitted for the Doctor's degree and deposited in the University of Kentucky Library are as a rule open for inspection, but are to be used only with due regard to the rights of the authors. Bibliographical references may be noted, but quotations or summaries of parts may be published only with the permission of the author, and with the usual scholarly acknowledgments.

Extensive copying or publication of the dissertation in whole or in part also requires the consent of the Dean of the Graduate School of the University of Kentucky.

A library that borrows this dissertation for use by its patrons is expected to secure the signature of each user.

Name Date

____

____

____

____

____

____

____

____

____

____

____

DISSERTATION

Charles Hokayem

The Graduate School
University of Kentucky

2010

ESSAYS ON HUMAN CAPITAL, HEALTH CAPITAL, AND THE LABOR MARKET

DISSERTATION

A dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy in the
College of Business and Economics
at the University of Kentucky

By

Charles Hokayem

Lexington, Kentucky

Director: Dr. James P. Ziliak, Carol Martin Gatton Chair in Microeconomics and

Director of the University of Kentucky Center for Poverty Research

Lexington, Kentucky

2010

ACKNOWLEDGEMENTS

During the completion of this dissertation, I have benefited from the insight and support of several people. I would like to thank Jim Ziliak for his time and intellectual discussions about this research. I cannot imagine a better dissertation advisor through this process. I would like to thank members of my dissertation committee, Glenn Blomquist, Chris Bollinger, and Julia Costich, for their helpful comments and advice. I am grateful to Bob Martin, my undergraduate advisor, for introducing me to economics my first year at Centre College. Most importantly, I thank my mother, father, and brother for always believing in me and encouraging me throughout graduate school.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 A STRUCTURAL MODEL OF LIFE CYCLE LABOR SUPPLY WITH ENDOGENOUS HEALTH AND HUMAN CAPITAL ACCUMULATION

1.1 Introduction

Labor economists have long documented the importance of human capital in labor market outcomes. Whether in the form of formal schooling or on-the-job training, human capital plays a crucial role in the development of wages and wealth over the life cycle (Card 1999). While education remains the most studied form of human capital, labor economists have studied human capital accumulation on the job through a learning-by-doing process (Ben-Porath 1967; Heckman 1976; Shaw 1989; Imai and Keane 2004). Learning-by-doing refers to the skills individuals acquire through working on the job separate from the skills acquired through formal schooling. As an individual spends more time on the job, more skills are acquired, leading to higher future income. Learning-by-doing models have generally ignored the role of health, or health capital, during this process. At the same time health economists have established the role of health in determining how much an individual works but have largely ignored the impact of health on acquiring skills on the job (Currie and Madrian 1999). If health determines how much an individual works and how much an individual works determines the skills acquired on the job, then variations in health will influence human capital accumulation and, subsequently, future income. This process suggests an interaction between health and human capital that has generally not been addressed in the health or labor literatures. This chapter combines these strands of literatures by introducing health in the spirit of Grossman (1972) to the standard learning-by-doing model.

This chapter develops a structural life cycle model to study the interaction between health and human capital before the retirement years. The basic idea is that a poor health shock has the direct effect of lowering labor supply in the current period which in turn reduces the amount of human capital gained on the job, resulting in an indirect effect of lower future wages. Set in a stochastic lifecycle framework, the model assumes an individual chooses hours worked, medical consumption, and nonmedical consumption to maximize lifetime utility subject to a budget constraint, a healthy time constraint, and production functions for human capital and health capital. The human and health capital production functions generate time nonseparabilities in the intertemporal budget constraints because stocks only depreciate gradually over time. The model assumes individuals face uncertainty about future realizations of health, tastes, prices, wages and interest rates. The forcing variable in the model is the assumption of exogenous health shocks that have a direct effect on sick time, which in turn affects optimal consumption and leisure choices. The estimation strategy follows a two step procedure. The first step estimates the human capital and health capital production parameters. The second step incorporates the production function parameters to estimate utility parameters using a nonlinear generalized method of moments (GMM) estimator derived from the solution to the individual's optimization problem.

To focus on the working years of the life cycle the model uses data on working male head of households from the Panel Study of Income Dynamics (PSID) who are between the ages of 25 and 60, spanning 1989-2003 waves. An innovation in the model is the incorporation of an often underutilized question about annual sick time into the standard time budget constraint, creating a budget constraint for healthy time. With this new healthy time budget constraint an individual can allocate healthy time to work or leisure. Since medical out-of-pocket expenditures are not always available in the PSID, the estimation adds the imputation of medical out-of-pocket expenditures from the Consumer Expenditure Survey as an input into the production of health.

The results show human capital production increases with the current stock of human capital at a decreasing rate which is in accord with diminishing returns to the human capital stock. Health capital production increases with the current stock of health capital, or better current health improves future health. Health has a relatively small role in interrupting the on-the-job human capital accumulation process. The model produces plausible estimates of the intertemporal substitution elasticity for consumption with the average between -1.51 and -0.88. Future work will address the measurement error in health and will also carry out counterfactual policy simulations. Simulations that compare the human capital and wage paths of individuals who experience different health shocks will illustrate the importance of good health. With the recent emphasis on health care reform simulations on subsidizing the price of medical care will provide insight into the effects of government health care policies on human and health capital accumulation. Initial predictions suggest a government policy subsidizing medical out-of-pocket expenditures causes individuals to have better health, accumulate more human capital, and earn higher wages in the long run.

1.2 Literature Review

1.2.1 Health and Labor Supply Literature

Grossman's (1972) model of health capital and the demand for health remains the most important theoretical contribution to health economics and serves as a framework for modeling health in this chapter. Grossman's model combines the idea of health as a capital stock with Becker's (1965) time allocation model. An individual inherits a stock of health that depreciates over time while purchasing medical care and devoting time to health investment (e.g. exercise) to replenish this stock. A time allocation constraint sums work time, sick time, time spent producing health, and time spent producing the household commodity. In this model the consumer maximizes a utility function that depends on the consumption of healthy days and a household commodity subject to a standard lifetime budget constraint which includes the purchase of medical care, labor income, and asset income. A consumer demands healthy days for two reasons: (1) consumption (sick days create disutility) and (2) investment (healthy days determine the time available for market and nonmarket activities and reduce time being sick). These two reasons for healthy days lead to a pure consumption model and a pure investment model. Human capital in this model is education and is taken as exogenous. Human capital differs from health capital in that human capital affects a consumer's productivity while health capital affects time spent being productive. Working with the pure investment version of this model Grossman derives health supply and demand functions whose intersection determines the optimal stock of health. The Grossman model forms the basis for reduced form estimation of the demand for health, demand for medical services, and health investment (Grossman 2000). The model also serves as motivation for empirical studies of health on labor supply to consider health as endogenous since it implies a conditional labor supply function that depends on the endogenous health stock (Currie and Madrian 1999). While the Grossman model provides a reason for health economists to study the relationship between health and education, it only treats human capital as exogenous and does not address human capital accumulation on the job (Grossman 2000; Grossman 2006).

Empirical testing of Grossman's model generally involves estimating reduced form demand equations for health, medical services, and health investment. The most common equations to estimate are demand for health and demand for medical services. These equations are functions of the model's exogenous variables: wage rate, price of medical care, the stock of human capital (education), and age. Grossman's investment model predicts the following signs: (1) demand for health equation: wage rate (+), price of medical care (-), human capital (+), age (-); (2) demand for medical services equation: wage rate (+), price of medical care (-), human capital (-), age(+). An increase in the wage implies a higher return to healthy time, so individuals demand for health and medical services increases with the wage. Standard microeconomic theory suggests the demand for health and medical services will fall with an increase in the price. Education improves the efficiency of investment in health, so more educated individuals require less inputs (e.g. medical services) to achieve a given level of investment. More educated individuals also receive a higher return for a given level of health, so they will demand higher levels of health. The depreciation rate of the health stock rises with age and raises the cost of holding the capital stock, so older individuals demand less health. The rising depreciation rate with age implies older individuals will also invest more in health by demanding more medical services. This feature of the model corroborates the observation that the elderly purchase more medical care as their health deteriorates. Grossman fits these equations to 1963 US data using either restricted activity days or work-loss days as a measure of healthy time and personal medical expenditures as a measure of medical care use. With no data on medical prices, Grossman's main independent variables are wage rate, family income, years of schooling, and age. Grossman finds the predicted signs on education, wage rate, and age in the health equation. However, he only finds the predicted sign of age in the medical services equation. The signs on wage and education are not consistent with the investment model.

Wagstaff (1986) tests the pure investment and pure consumption versions of the Grossman model using the 1976 Danish Welfare Study. Realizing the multidimensional nature of health Wagstaff uses the multiple indicators-multiple causes technique (MIMIC) to condense nineteen different health measures into a single measure of unobserved health stock. Wagstaff calls a version of his model with assets and lifetime wage rates a pure investment model and a version without assets and lifetime wage rates a pure consumption model. His estimation of the health demand function in the pure investment model produces the correct signs on wage rate, schooling, and age. His medical care demand function with physician visits produces a negative coefficient on schooling which, unlike Grossman's results, is consistent with the pure investment model. Erbsland, Ried, and Ulrich (1995) also use the MIMIC technique but with 1986 West German Socio-Economic Panel data. They rely on four self-reported indicators of the unobserved stock of health. Their results are consistent with Grossman's pure investment model. In their health demand function schooling is positive while age is negative. In their medical services demand function based on general practitioner visits, age is positive and schooling is negative.

Sickles and Yazbeck (1998) estimate a structural model of life cycle leisure demand with health production. They focus on the impact of health-related spending and variations in leisure demand on the production of good health for the older population. Their model incorporates temporal nonseparability in leisure, health, and consumption along with intertemporal nonseparability in health. A distributed lag of past health affects current health accumulation in the same way as Hotz, Kydland, and Sedlacek (1988) allow intertemporal nonseparability of leisure. Individuals in their model choose leisure, health-related consumption, and health-neutral consumption. Health production depends on health-related consumption and leisure. They estimate their model using nonlinear generalized method of moments on a panel of older men (age 58-63) from the Retirement History Survey 1969-1979. The Quality Well-Being Index serves as a continuous measure of health. It combines scales measuring mobility, physical activity, social activity, and self-assessed health status. Overall, Sickles and Yazbeck estimate positive health-related consumption and leisure elasticities, suggesting health-related consumption and leisure lead to better health. The health-related consumption elasticity (change in health due to change in health-related consumption) is estimated around .03, so a 10 percent increase in health consumption leads to about a 0.3 percent improvement in health. This small effect supports the finding of "flat-of-the-curve" medicine in the health literature.[1] The corresponding leisure elasticity ranges from .23 to .69, so a 10 percent increase in leisure leads to a 2.3 to 6.9 percent improvement in health. Sickles and Yazbeck find individuals have a short memory of past health on current utility, and past health is important to current health, implying a role for dynamics in the health production process.

[1] "Flat-of-the-curve" medicine refers to the small marginal impact of medical services on health status.

Gilleskie (2010) studies an individual's daily decision to miss work and/or visit the doctor during an acute illness episode. Gilleskie estimates a discrete choice structural model with the goal of understanding the differences between men and women when making decisions about being absent from work and seeking medical care *during* an episode of illness. The model allows for the presence of a sick leave policy that replaces lost income and a health insurance policy that provides varying coinsurance rates. Individuals face a probability of contracting an illness that depends on health status and a probability of recovering from an illness that depends on work absences, physician visits, and length of current illness. Medical care occurs only when an individual is sick. Seeking preventative treatment is not allowed. Gilleskie estimates the model on a sample of employed men and women age 25-64 from the 1987 National Medical Expenditure Survey. Individuals in the sample have illnesses that last less than 22 days. Examples of illnesses include the common cold, the flu, strep throat, and bronchitis. Gilleskie's results suggest sickness creates disutility for men and women. An additional dollar of income provides less utility when sick than well for men and women. Absences are more productive than medical care as they raise the probability of recovering from an illness by as much as 3 percent. Absences and seeking treatment are complements in recovering from an illness. Individuals in poorer health recover slower from an illness. Women are more likely than men to suffer from an illness for which they choose to miss work or seek medical treatment.

Gilleskie uses her structural model to simulate several policy experiments that change the availability of sick leave, health insurance, and the price of medical care. Overall, the policy experiments highlight the differences in responses between men and women. Changing from no sick leave coverage to sick leave coverage induces men to take more absences (45 percent increase) than women (11 percent increase). In a policy providing physician visits at no cost compared to a policy where individuals pay the full price of care men increase their medical consumption by 20 percent per illness episode while women increase their medical consumption by 15 percent.

1.2.2 Health and Labor Supply

Most of the labor literature addresses the effect of health on static labor supply with little attention on health and life cycle labor supply, the exception being research that examines health around the age of retirement (Currie and Madrian 1999; French 2005; Bound, Stinebrickner, and Waidmann 2010). This literature highlights two important issues with health and labor supply: the treatment of health endogeneity and the measurement of health.

Currie and Madrian (1999) extensively review the labor literature connecting health to labor market behavior. Their review focuses on the impact of health on wages, earnings, hours worked, labor force participation, and type of work. There seems to be no work on the effect of health on human capital accumulation and little work on life cycle labor supply. Two estimation issues in this literature are the endogeneity of health and the measurement error in health. Motivated by the Grossman model they emphasize the need to consider health endogenous. They also highlight the endogeneity associated with measurement error in health. Often used self-reported health measures suffer from endogeneity in that healthier individuals tend to work more. Self-reported measures are also problematic due to nonrandom measurement error. Individuals who don't work or cut down their hours worked could be more likely to report poor health to justify their labor market behavior. Individuals could also misreport their health to stay in government assistance programs. Self-reported measures may also be affected by whether the individual has sought treatment. A large part of the literature has documented this nonrandom measurement error.

From their review Currie and Madrian (1999) draw a few conclusions about hours worked and labor force participation. Overall, they note that studies tend to find a larger positive effect on hours worked than wages. Studies treating health as endogenous tend to find a smaller effect of health on hours worked than studies that treat health as exogenous. Better health also improves labor force participation, but there is wide disagreement on the size of this effect. The estimated effects of health on labor market outcomes vary and are sensitive to identification assumptions and the way health is measured.[2] The samples in the literature tend to focus on older white men.[3] They acknowledge the need to take a structural approach to connecting health to labor market behavior with the use of a health production function like Sickles and Yazbeck (1998). They also suggest future work should use more general samples.

Hum, Simpson, and Fissuh (2006) and Hum, Simpson, and Fissuh (2008) directly examine health on life cycle labor supply and employment with panel data on Canadaian men during the working years of life. Hum, Simpson, and Fissuh add a health stock to the

---

[2] Health measures can fall into the following categories: (1) self-reported health; (2) health limitations on the ability to work; (3) functional limitations such as problems with activities of daily living (ADL's); (4) presence of chronic or acute conditions; (5) utilization of medical care; (6) clinical assessment of a condition; (7) nutritional status; and (8) expected or future mortality.

[3] Two exceptions are Berger (1983) and Berger and Fleisher (1984) which examine spousal health and labor supply using couples during the working years of life.

standard life cycle labor supply model and derive employment and labor supply functions which they estimate using logit and tobit models, respectively. They define employment as having hours worked greater than zero. Both specifications address the endogeneity and mismeasurement of health, two concerns raised by Currie and Madrian (1999), using an approach developed by Bound, Schoenbaum, Stinebrickner, and Waidmann (1999). The approach uses exogenous demographic variables (age, education, mother's education) and more objective health measures (functional limitations and work limiting functional limitations) to instrument for a self-reported rating of personal health (excellent, very good, good, fair, and poor). The instrumenting equation is either a probit model or an ordered probit model. The probit model covers good health and poor health with good health including excellent, very good, or good ratings and poor health including fair or poor rating. The ordered probit covers all health ratings. Both employment and labor supply specifications also allow for unobserved heterogeneity through fixed effects and control for sample selection bias due to unobserved wages.

Hum, Simpson, and Fissuh (2006) and Hum, Simpson, and Fissuh (2008) fit these equations with panel data of men age 21-65 from the Canadian Survey of Labor and Income Dynamics (SLID) for 1996-2001. The authors claim the Canadian data offer the advantage of studying health effects on labor supply in the context of universal health coverage which avoids any confounding price effects due to varying health insurance coverage that would occur in the United States.[4] Overall, the effect of health on hours worked or employment is smaller when health is treated as endogenous and even smaller when using fixed effects. For the pooled sample and exogenous health the estimated

[4] While Hum, Simpson, and Fissuh claim there are no confounding price effects, they ignore differences in waiting times and access that can affect the price of care.

marginal effect of good health is a labor supply increase of 783 hours per year (43.5% of total hours worked by all men). Treating health as endogenous produces an estimated marginal effect of 486 hours (27.0% of total hours worked by all men). When including fixed effects, the estimated marginal effect falls to 119 hours and 91 hours for exogenous and endogenous health, respectively. Comparing these estimates suggests the bias from unobserved heterogeneity is larger than the bias from endogenous health. Unlike the previous literature that focuses on older men, Hum, Simpson, and Fissuh (2008) analyze subsamples by age group. The impact of health on hours worked is significant for all age groups (age 25-34, age 35-50, age 51-65) not just older men as previously studied. The effect increases across age groups with older men experiencing the largest marginal effect (marginal effect of good health: age 25-34: 4.2%; age 35-50: 4.3%; age 51-65: 6.0%).

1.2.3 Human Capital and Labor Supply Literature

The learning-by-doing model in the labor literature extends the standard life cycle labor supply model by allowing human capital investment through a learning-by-doing process. Learning-by-doing builds on the work of Ben-Porath (1967), Weiss (1972), and Heckman (1976). Human capital investment occurs as a by-product of market work. More hours on the job leads to more human capital. Unlike the standard model where the wage is exogenous, the wage in the learning-by-doing model depends on the level of human capital stock and the rental rate of the human capital. Since the human capital stock depends on hours worked, past labor supply decisions can affect the wage. This model adds an additional trade-off: the individual must consider the increase in utility by reducing current hours worked against the increase in future wage that occurs from learning on the job.

Shaw (1989) estimates a structural learning-by-doing model using a sample of men age 18-64 from the Panel Study of Income Dynamics (PSID) for the years 1967-1980. Shaw formulates the individual's problem in a dynamic programming framework and derives an Euler equation for the consumption and leisure choice. Her estimation strategy follows a two-step process. In the first step she estimates a wage equation by nonlinear instrumental variables to recover the unobserved rental rate on human capital and parameters of the quadratic human capital production function. The second step incorporates the estimated rental rates and human capital production parameters into the estimation of the Euler equation by generalized method of moments. The second step exploits orthogonality conditions from rational expectations to estimate parameters of the translog utility function.[5] Shaw finds strong evidence in favor of endogenous wages, that is, current hours worked positively affect future wages. A temporary 25 percent increase in hours of work increases the following year's wages by 12.8 percent. In addition, she presents simulation results showing the intertemporal elasticity of substitution for labor supply is not constant across the life cycle; instead, it rises over the life cycle.[6] An increasing intertemporal elasticity of substitution implies younger workers are less responsive to wage or tax changes than older workers.

Imai and Keane (2004) extend Shaw (1989) by using a more computationally intensive estimation procedure to estimate a similar model. Their technique uses a numerical dynamic programming technique to solve the individual's optimization

---

[5] Shaw's estimation and identification method is similar to Hansen and Singleton's (1982) generalized method of moments estimation of consumption Euler equations.

[6] The intertemporal elasticity of substitution for labor supply describes the percent change in hours worked due to a percent change in the wage.

problem and embed this solution in a simulated maximum likelihood estimation. This procedure allows for measurement error in wages, hours worked, and assets. Unlike Shaw who does not specifically estimate the intertemporal elasticity of substitution, Imai and Keane choose a utility function that identifies a point estimate. Their data, a sample of white males from the National Longitudinal Survey of Youth (NLSY), produce an intertemporal elasticity of substitution estimate of 3.82. This estimate suggests a 1 percent increase in the wage leads to a nearly 4 percent increase in hours worked. This estimate far exceeds estimates from the microliterature which range from 0.37 to 0.88 (MaCurdy 1981; Altonji 1986). Imai and Keane attribute the stark difference to previous studies omitting human capital accumulation.

Shaw (1989) and Imai and Keane (2004) do not address the role of health in learning-by-doing models. These studies provide a theoretical framework for incorporating health into a labor supply model with human capital. Their estimation methods give a useful starting point for developing an estimation and identification strategy to estimate structural parameters for a model that includes health. This model is described in the following sections.

1.3 Model of Health, Human Capital Accumulation, and Life Cycle Labor Supply

The model draws on elements from the health and labor literatures by combining health capital with human capital through a learning-by-doing technology. It allows health to enter utility directly and act as a constraint on behavior. An individual maximizes a standard utility function that is additively separable over time and is defined over leisure (L_t), nonmedical consumption (C_t), and the stock of health (H_t). In each period,

$$U = U(L_t, C_t, H_t)$$

Income comes from two sources, asset income ($r_t A_t$) and labor income ($w_t N_t$). This income can be spent on nonmedical consumption and used to purchase medical services (M_t) at the price p_t^m. Putting income and expenditures together gives the following intertemporal budget constraint:

$$A_{t+1} = (1 + r_t)(A_t + w_t N_t - C_t - p_t^m M_t) \text{ (Asset Accumulation)}$$

The observed wage (w_t) is the product of a human capital stock (K_t) and the unobserved rental rate on human capital (R_t), so unlike the standard labor supply model the wage is endogenous.

$$w_t = R_t K_t \text{ (Wage Equation)}$$

In each period an individual inherits a stock of human capital which depreciates at rate δ_K. Human capital evolves in a similar way as health capital. Next period's human capital is the previous period's stock less depreciation plus new investment. New investment occurs on the job through learning-by-doing that depends on hours worked and the levels of human and health capital, $x(N_t, K_t, H_t)$. Human capital evolves according to

$$K_{t+1} = (1 - \delta_K)K_t + x(N_t, K_t, H_t) = f(N_t, K_t, H_t) \text{ (Human Capital Accumulation)}$$

In each period an individual inherits a stock of health capital (H_t) which depreciates at rate δ_H and can be replenished by devoting time to health, L_t (e.g.

exercise), and purchasing medical services. Leisure and medical services enter a health production function, $y(M_t, L_t)$.[7] Health capital then evolves according to

$$H_{t+1} = (1 - \delta_H)H_t + y(M_t, L_t) = I(M_t, L_t, H_t) \text{(Health Capital Accumulation)}$$

Individuals also have healthy time (ht_t) which can be allocated across two activities: leisure (L_t) and work (N_t). The healthy time constraint becomes

$$L_t + N_t = ht_t \text{ (Healthy Time Constraint)}$$

In addition, total time (*T*) is the sum of healthy time (ht_t) and sick time (s_t), so

$$ht_t + s_t = T \text{ (Total Time Constraint)}$$

The individual's optimization problem can be represented in a dynamic programming framework with state variables, (A_t, K_t, H_t), and choice variables (C_t, L_t, M_t).

$$V(A_t, K_t, H_t) = \max_{C_t, L_t, M_t} \{U(L_t, C_t, H_t) + \beta V(A_{t+1}, K_{t+1}, H_{t+1})\}$$

$$s.t.\, A_{t+1} = (1 + r_t)(A_t + w_t N_t - C_t - p_t^m M_t) \text{ (Asset Accumulation)}$$

$$w_t = R_t K_t \text{ (Wage Equation)}$$

$$K_{t+1} = (1 - \delta_K)K_t + x(N_t, K_t, H_t)$$

$$= f(N_t, K_t, H_t) \text{ (Human Capital Accumulation)}$$

$$H_{t+1} = (1 - \delta_H)H_t + y(M_t, L_t) = I(M_t, L_t, H_t) \text{(Health Capital Accumulation)}$$

[7] The concept of leisure in this framework differs from leisure in the standard labor supply model where leisure represents nonmarket time. The Grossman (1972) model separates total time into time for work, time for health, time for producing the household good, and time for sickness. The ideal set of data would consist of time divided into these categories. Time diary data are not available for the years used in the estimation. Leisure in this framework is meant to capture time input into health production.

$L_t + N_t = ht_t$ (Healthy Time Constraint)

$ht_t + s_t = T$ (Total Time Constraint)

Like other life cycle labor supply models the source of uncertainty in this model comes from unknown future realizations of tastes, prices, wages, and interest rates (MaCurdy 1983). This model does not make assumptions about the form of the distributions generating the uncertainty. Unlike other models of life cycle labor supply, this model adds uncertain health shocks through sick time. s_t also follows a distribution of unknown form. The individual in this model does not know when she will be sick. s_t characterizes acute illnesses that an individual cannot anticipate such as the common cold, food poisoning, and the flu.[8] s_t acts as an exogenous shock to the total time and healthy time constraints.[9]

The optimization problem can be simplified by substituting the wage equation and time constraints. With choice variables $C_{t,}$, N_t, and M_t the optimization problem can now be written as

$$V(A_t, K_t, H_t) = \max_{C_t, N_t, M_t} \{U(ht_t - N_t, C_t, H_t) + \beta V(A_{t+1}, K_{t+1}, H_{t+1})\}$$

$$s.t.\, A_{t+1} = (1 + r_t)(A_t + R_t K_t N_t - C_t - p_t^m M_t) \text{ (Asset Accumulation)}$$

$$K_{t+1} = f(N_t, K_t, H_t) \text{ (Human Capital Accumulation)}$$

$$H_{t+1} = I(M_t, ht_t - N_t, H_t) \text{(Health Capital Accumulation)}$$

The first order conditions with respect to C_t, N_t, and M_t:

[8] Gilleskie (2010) also considers acute illnesses.

[9] This idea of s_t affecting the time constraint is similar in spirit to the literature on fixed time costs and labor supply (Cogan 1981).

$$C_t\colon U_{c,t} - \beta(1+r_t)E_t\{V_A^{t+1}\} = 0 \quad (1)$$

$$N_t\colon - U_{L,t} + \beta E_t\{V_A^{t+1}(1+r_t)R_t K_t + V_K^{t+1} f_{N,t} - V_H^{t+1} I_{L,t}\} = 0 \quad (2)$$

$$M_t\colon \beta E_t\{-V_A^{t+1}(1+r_t)p_t^m + V_H^{t+1} I_{M,t}\} = 0 \quad (3)$$

A subscript represents a partial derivative with respect to that variable at time *t*. For example, $U_{c,t} = \frac{\partial U_t}{\partial C_t}$ and $U_{c,t+1} = \frac{\partial U_{t+1}}{\partial C_{t+1}}$. Equations 1 and 2 are most similar to conditions in life cycle labor supply models. Equation 1 is the standard Euler equation for consumption and describes the optimal consumption over time. Equation 2 shows the effect of endogenous human capital and health capital accumulation in the last two terms. Without endogenous human capital and health capital accumulation, these terms are zero, and Equation 2 reduces to a condition similar to the standard labor-leisure condition from the static labor supply model where $R_t K_t$ is replaced by w_t. Equation 3 describes the optimal amount of medical consumption. The first order conditions can be combined with envelope conditions for the state variables (A_t, K_t, H_t) to solve the optimization problem. The following equilibrium condition characterizes the solution to the optimization problem.[10]

$$-U_{L,t} + \beta U_{c,t+1}(1+r_t)w_t + \beta f_{N,t} U_{c,t+1} R_{t+1} N_{t+1} - \frac{U_{c,t} p_t^m}{I_{M,t}} I_{L,t} +$$

$$\beta f_{N,t} \frac{f_{k,t+1}}{f_{N,t+1}} \left(U_{L,t+1} - U_{c,t+1} w_{t+1} + \frac{U_{c,t+1} p_{t+1}^m}{I_{M,t+1}} I_{L,t+1} \right) = 0 \quad (4)$$

[10] See Appendix for derivation of solution to the optimization problem.

Under rational expectations, eq(4) has zero expectation at time period t, so realizations of future variables imply

$$-U_{L,t} + \beta U_{c,t+1}(1 + r_t)w_t + \beta f_{N,t} U_{c,t+1} R_{t+1} N_{t+1} - \frac{U_{c,t} p_t^m}{I_{M,t}} I_{L,t} +$$

$$\beta f_{N,t} \frac{f_{k,t+1}}{f_{N,t+1}} \left(U_{L,t+1} - U_{c,t+1} w_{t+1} + \frac{U_{c,t+1} p_{t+1}^m}{I_{M,t+1}} I_{L,t+1} \right) = u_{t+1} \text{ (5)}$$

where u_{t+1} is the forecast error at time *t*. Rational expectations implies $E_t\{u_{t+1}\} = 0$, so any information at time *t* is not useful in forecasting future variables. This orthogonality between u_{t+1} and the information at time *t* will be exploited to estimate the structural parameters of the model.

1.4 Econometric Specification

A two step estimation strategy is used to estimate the structural parameters of the model. In the first step quadratic human capital and health production function parameters are estimated with ordinary least squares. The second step incorporates these production function parameters in a nonlinear Generalized Method of Moments (GMM) estimation of the equilibrium condition (5) to identify utility parameters.

The specification for the human capital production function, $f(N_t, K_t, H_t)$, is quadratic in its arguments. The quadratic specification represents the concave nature of earnings over the life-cycle:

$$K_{t+1} = f(N_t, K_t, H_t) = \alpha_0 K_t + \alpha_1 K_t^2 + \alpha_2 K_t N_t + \alpha_3 N_t + \alpha_4 N_t^2 +$$

$$\alpha_5 H_t + \alpha_6 H_t^2 + \alpha_7 H_t N_t + \alpha_8 K_t H_t + \tau_t + \epsilon_{i,t}$$

The vector $\alpha = (\alpha_0, \alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6, \alpha_7, \alpha_8)$ contains the structural parameters of the human capital production function. τ_t is an exogenous time-specific growth rate of human capital common to all individuals, and $\epsilon_{i,t}$ is an individual-specific component of human capital growth. This specification differs from Shaw (1989) in its inclusion of health into the human capital production function. As written, the human capital production function can not be directly estimated since the stock of human capital, K_t, is not observed. Wages are observed, so the specification can be rewritten in terms of wages using the relationship, $w_t = R_t K_t$. Making the substitution $K_t = \frac{w_t}{R_t}$ gives a specification that can be estimated

$$\frac{w_{t+1}}{R_{t+1}} = \alpha_0 \frac{w_t}{R_t} + \alpha_1 \left(\frac{w_t}{R_t}\right)^2 + \alpha_2 \left(\frac{w_t}{R_t}\right) N_t + \alpha_3 N_t + \alpha_4 N_t^2 +$$

$$\alpha_5 H_t + \alpha_6 H_t^2 + \alpha_7 H_t N_t + \alpha_8 \left(\frac{w_t}{R_t}\right) H_t + \tau_t + \epsilon_{i,t}$$

This specification of the wage equation relates future discounted wages to current hours worked, current discounted wages, and current health. It is estimated using ordinary least squares, setting $R_t = 1\ \forall t$.[11] This specification of a dynamic wage equation with hours worked is similar to the standard Mincer equation. The Mincer equation is derived from a lifecycle earnings model with a human capital stock where an individual maximizes the present discounted value of earnings. The equation is a quadratic function of one's labor market experience, typically approximated by years of potential experience (Card 1999; Heckman, Lochner, and Todd 2006; Polachek 2008). The learning-by-doing model represents an alternative approach where an individual

[11] Attempts to estimate annual rental rates were unsuccessful.

maximizes lifetime utility instead of wealth and incorporates the traditional labor leisure supply model into the human capital acquisition process which relies on hours worked (Polachek 2007). The Mincer equation has been estimated with years of tenure on-the-job in place of years of potential experience. Tenure on-the-job would represent an approximation to hours worked.

The marginal product of hours worked ($f_{N,t}$), marginal product of human capital stock ($f_{K,t}$), and the marginal product of health capital stock ($f_{H,t}$) come from differentiating the quadratic human capital production function

$$f_{N,t} = \alpha_2 \left(\frac{w_t}{R_t}\right) + \alpha_3 + 2\alpha_4 N_t + \alpha_7 H_t$$

$$f_{K,t} = \alpha_0 + 2\alpha_1 \frac{w_t}{R_t} + \alpha_2 N_t + \alpha_8 H_t$$

$$f_{H,t} = \alpha_5 + 2\alpha_6 H_t + \alpha_7 N_t + \alpha_8 \frac{w_t}{R_t}$$

Health capital production, $I(M_t, L_t, H_t)$, follows a quadratic specification similar to human capital production. The health capital production is specified as

$$H_{t+1} = I(M_t, L_t, H_t) = \theta_0 H_t + \theta_1 H_t^2 + \theta_2 H_t L_t + \theta_3 L_t + \theta_4 L_t^2 +$$

$$\theta_5 M_t + \theta_6 M_t^2 + \theta_7 H_t M_t + \theta_8 L_t M_t + \tau_t + v_{i,t}$$

where the vector $\theta = (\theta_0, \theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6, \theta_7, \theta_8)$ represents the structural parameters of health capital production. τ_t is an exogenous time-specific growth rate of health capital common to all individuals, and $v_{i,t}$ is an individual-specific component of health capital growth. The corresponding marginal product of medical expenditures ($I_{M,t}$), leisure ($I_{L,t}$), and health stock ($I_{H,t}$) are

$$I_{M,t} = \theta_5 + 2\theta_6 M_t + \theta_7 H_t + \theta_8 L_t$$

$$I_{L,t} = \theta_2 H_t + \theta_3 + 2\theta_4 L_t + \theta_8 M_t$$

$$I_{H,t} = \theta_0 + 2\theta_1 H_t + \theta_2 L_t + \theta_7 M_t$$

The specification of the utility function follows a translog form. The translog utility function is quadratic in its arguments and represents a local second-order approximation to any utility function. It also does not impose the restrictions of additivity or homothecity associated with other common utility functions, such as CES or Cobb-Douglas (Christensen, Jorgenson, and Lau 1975). The exact specification is

$$U(L_t, C_t, H_t) = \gamma_0 \ln L_t + \gamma_1 \ln C_t + \gamma_2 (\ln L_t)(\ln C_t) + \gamma_3 (\ln L_t)^2 + \gamma_4 (\ln C_t)^2 +$$

$$\gamma_5 \ln H_t + \gamma_6 (\ln H_t)(\ln L_t) + \gamma_7 (\ln H_t)(\ln C_t) + \gamma_8 (\ln H_t)^2$$

The vector $\gamma = (\gamma_0, \gamma_1, \gamma_2, \gamma_3, \gamma_4, \gamma_5, \gamma_6, \gamma_7, \gamma_8)$ represents the structural parameters of the utility function. Identification of these parameters requires a normalization, so $\gamma_0 = 1$. Due to the form of the equilibrium condition (5) γ_5 and γ_8 can not be identified so are set to zero. Differentiating the translog utility function gives the marginal utility of leisure ($U_{L,t}$) and the marginal utility of nonmedical consumption ($U_{C,t}$) as

$$U_{L,t} = \frac{1 + \gamma_2 \ln C_t + 2\gamma_3 \ln L_t + \gamma_6 \ln H_t}{L_t}$$

$$U_{C,t} = \frac{\gamma_1 + \gamma_2 \ln L_t + 2\gamma_4 \ln C_t + \gamma_7 \ln H_t}{C_t}$$

Substituting the corresponding marginal utility and marginal product functions into the equilibrium condition produces a highly nonlinear equation that can be parameterized by $\Gamma = (\alpha, \gamma, \theta, \beta)$. Γ is the vector of unknown population parameters. α and θ are estimated from the human capital and health capital production functions in the first step, and β is set to .95, leaving the utility parameters to be estimated in the second step.[12] Let X_{it} be the vector of variables entering the *i*th individual's equilibrium condition in period *t*. The *i*th individual's equation can be represented by

$$f(X_{it}, \gamma; \alpha, \theta, \beta) = u_{it+1}$$

Rational expectations implies information in the information set Ω_{it} is not useful in forecasting future variables, so

$$E_t\{f(X_{it}, \gamma; \alpha, \theta, \beta) \cdot Z_{it}\} = 0$$

where Z_{it} contains elements of Ω_{it}. The orthogonality between $f(X_{it}, \gamma; \alpha, \theta, \beta)$ and Z_{it} is exploited to estimate γ in a GMM estimator. With panel data of *T* years for each individual population orthogonality conditions are derived by averaging over time,

$$E\frac{1}{T}\sum_{t=1}^{T}[f(X_{it}, \gamma; \alpha, \theta, \beta)Z_{it}] = E[M(X_i, Z_i, \gamma; \alpha, \theta, \beta)] = 0$$

The sample analogs of these population conditions are constructed by averaging over a random sample of *N* individuals, so

[12] The production and utility parameters can be estimated at one time as a system. Attempts to estimate the system were computationally infeasible.

$$O_N(\gamma) = \frac{1}{N}\sum_{i=1}^{N}[M(X_i, Z_i, \gamma; \alpha, \theta, \beta)]$$

The GMM estimator of γ minimizes the quadratic form $O_N(\gamma)W_N O_N'(\gamma)$ or

$$\gamma = \underset{\gamma}{\operatorname{argmin}}\, O_N(\gamma)W_N O_N'(\gamma)$$

where the W_N is a symmetric positive definite weighting matrix. W_N is unknown so is replaced by $\widehat{W_N}$ constructed from the residuals of a nonlinear two stage least squares (NL2SLS) procedure, allowing for conditional heteroskedasticity.

1.5 Data Construction

The Panel Study of Income Dynamics (PSID) serves as the main data source for estimating the structural model. The PSID offers the advantage of covering a representative sample of U.S. individuals during the working portion of the life cycle. The PSID started with 4,800 families in 1968 and with efforts to follow all family members now contains over 7,000 families. Interviews occur annually for 1968-1996 and biennially for 1997-2007. The PSID consists of three samples: the original Survey Research Center (SRC) sample, the Survey of Economic Opportunity (SEO) sample, and the Latino sample. The original SRC sample represents the U.S. population while the SEO sample serves as a supplementary low income subsample. The Latino sample, added in 1990, accounts for the changing nature of immigration in the U.S. This chapter excludes the SEO and Latino samples. Individuals are drawn from the SRC sample from 1989-2003 who meet the following criteria: (1) in family at time of interview; (2) head of household; (2) male; (3) not self-employed; (4) employed; (5)work at least 3 years; (6)

age between 25 and 60; (7) real annual food consumption between $520 and one-third of real annual family income; (8) annual hours worked at least 100; and (9) real hourly wage between $2/hour and $200/hour. Limiting the sample to male head of households of working age avoids issues with labor force nonparticipation and joint labor supply decisions. The sample selection criteria produce an unbalanced panel. This analysis treats the panel as continuous for the period 1989-2003 (ignoring biennial interruptions) and considers missing person-year observations as missing conditionally at random.

The PSID collects a variety of information on the labor market and socioeconomic characteristics of each household. Questions about annual family income and annual hours worked refer to the previous calendar year.[13] Food consumption also refers to the previous calendar year. Total annual food consumption includes food at home, food delivered to the home, eating out, and the value of food stamps. Total food consumption, deflated by the food component of the Consumer Price Index (2008 base year), is one of the consumption measures in the model. The PSID collects food consumption each survey year with the exception of survey years 1988 and 1989. For survey years after 1993, the PSID stopped reporting annual values for food at home, food delivered, eating out, and the value of food stamps. For these years annual amounts were created by annualizing the amount reported according to the reporting period (daily, weekly, biweekly, monthly, and monthly). An alternative consumption measure in the model is a nondurable consumption measure imputed from food consumption reported in the PSID. The appendix describes the imputation procedure.

[13] Annual family income with negative values after 1993 is set to 1 to match the PSID bottom coding of family income in previous years.

The hourly wage rate depends on whether the worker is hourly or salaried. For hourly workers the PSID gives the reported wage rate. For salaried workers the PSID reports an hourly rate that depends on the salary and the pay period (weekly, biweekly, or monthly). For these workers the PSID constructs a wage adjusted for a fixed number of hours each pay period instead of using the actual hours worked. For example, a salaried worker paid weekly has an hourly rate that is the salary divided by 40 hours. Salaried workers paid biweekly, monthly, and yearly have salaries divided by 80, 160, and 2,000 hours, respectively. All wage and income data are deflated by the personal consumption deflator using a 2008 base year.

Questions about annual sick hours and self-rated health serve as measures of healthy time and the health stock, respectively. PSID respondents report the amount of work missed due to own illness for the previous year. Specifically, the PSID asks "Did you miss work because you were sick? How much work did you miss?" The PSID calculates the annual sick hours as weeks ill times 8 hours for the first 8 weeks and times 60 hours for any weeks thereafter. Beginning in 1994 the PSID did not calculate annual sick hours, so annual sick hours are created by applying the same formula to the number of reported days, weeks, or months missed due to own sickness. The PSID self-rated health question began in 1984 and asks "Would you say your health in general is excellent, very good, good, fair, or poor?"[14] These responses are converted to a four point scale in the following way: 1 (Fair or Poor); 2 Good; 3 Very Good; and 4 Excellent.

[14] The self-rated health question in the PSID has been used in the recent health literature. Fletcher and Sindelar (2009) and Fletcher, Sindelar, and Yamaguchi (2009) are examples.

Other variables necessary for the model are the interest rate and medical out-of-pocket expenditures. The interest rate in the model reflects the after-tax annual 3-month Treasury bill interest rate based on a marginal federal tax rate from the NBER TAXSIM module.[15] Medical out-of-pocket expenditures are imputed using data from the Consumer Expenditure Survey. The appendix explains the imputation procedure for medical out-of-pocket expenditures.

After imputing consumption and health expenditure data and applying the sample selection criteria, the analysis sample contains 15,236 person-year observations and covers 1989-2003. Table 1.1 displays summary statistics. The sample consists of males, who are mostly white, have completed almost 14 years of education, and on average are 39 years old. These men earn an average hourly wage of $25.15/hour and work about 2,300 hours each year. With an average health status just under 3, health in the sample is about "Very Good." Individuals take just over 35 hours of sick time each year and spend almost $2,900 on medical out-of-pocket expenditures. Figure 1.1 shows the distribution of sick hours for the sample. A large number of individuals are concentrated at zero sick hours. Sick hours at the 25th and median percentiles are zero while sick hours at the 75th percentile are 32. The second panel of Table 1.1 shows summary statistics for the sample by sick hours (sick hours equal to zero, sick hours between zero and the 75th percentile, and sick hours greater than the 75th percentile). Men with sick hours between zero and the 75th percentile have the highest average hourly wage ($25.84), are in the best health (3.14), and spend the most on medical out-of-pocket expenditures ($2947). Average health and work hours are the lowest among men with sick hours greater than the 75th

[15] The marginal federal tax rate comes from the NBER TAXSIM module using head of household, labor income, and number of children as input variables.

percentile (health statuts of 2.90, work hours of 2209). Several variables must be scaled to facilitate the computation of the model. Leisure, hours worked, and healthy time are divided by 1,000. Total food consumption and medical out-of-pocket expenditures are divided by 10,000. The hourly wage and age are divided by 100.

1.6 Results

Table 1.2 and Table 1.3 show the estimation of the human capital and health capital production functions using ordinary least squares. Heteroskedasticity robust standard errors are in parentheses. Columns 1 and 2 of each table provide parameter estimates for the base specifications while columns 3 and 4 introduce education and demographic heterogeneity. This interaction with education and nonwhite allows education and nonwhite to shift the production functions. Both human capital specifications clearly show a concave relationship between current wages and future wages. Assuming the wage represents the human capital stock, $K_t = \frac{w_t}{R_t}$, the results show human capital production increases with the current stock of human capital at a decreasing rate, suggesting decreasing marginal productivity of the human capital stock. Education augments the future wage through the quadratic current wage. Overall, the insignificant parameter estimates for health suggest health plays a relatively weak role in interrupting the human capital accumulation process. The specification without heterogeneity suggests for a given level of human capital health improves human capital accumulation. Both health capital specifications indicate better current health improves future health. The specification without heterogeneity shows for a given level of health, more out-of-pocket medical expenditures improves future health. Both specifications show for a given level of health more time devoted to leisure also improves future health.[16] The effect of current health on future health is larger for nonwhite men. The effect of out-of-pocket medical expenditures on future health is smaller for nonwhite men. Nonwhite men experience differing quadratic effects of current health (negative) and out-of-pocket medical expenditures (positive) on future health.

Tables 1.4-1.9 show the nonlinear generalized method of moments estimation of utility parameters.[17] These tables follow the same layout as Table 1.2 and Table 1.3. Columns 1 and 2 are based on the estimated production parameters without heterogeneity while columns 3 and 4 are based on production parameters that allow heterogeneity. The coefficient on leisure is set to 1 for identification of the remaining parameters. The tables differ in the timing of the instrument set. Table 1.4 and Table 1.7 use contemporaneous instruments; Tables 1.5 and Table 1.8 add (*t-1*) instruments; Table 1.6 and Table 1.9 add (*t-2*) instruments. Tables 1.7-1.9 report utility parameters based on nondurable consumption and contemporaneous instruments. The instrument set used in the estimation includes leisure, food consumption, wage, health status, medical expenditures, the after-tax interest rate, age, education, region dummies, nonwhite, number of children, family size, annual time dummies, interactions between leisure and food consumption, wage and health status, health status and leisure, medical expenditures and health status, leisure and medical expenditures, and age and education. Squared values of food consumption, wage, health status, medical expenditures, age, and leisure are also included.

[16] This result is broadly consistent with Ruhm (2000) who finds recessions are good for your health. As the economy enters a downturn, individuals work less hours and have more leisure time, making it less costly to pursue time intensive health production activities such as exercise.
[17] The nonlinear optimization uses the Newton-Raphson method.

Table 1.4 shows leisure exhibits strong diminishing returns with and without heterogeneity which agrees with Shaw (1989) who estimates a similar model without health effects. Introducing heterogeneity produces a positive coefficient on the interaction with leisure and consumption, suggesting they are complements.
When evaluated at the sample means, the marginal utility of consumption is negative with and without heterogeneity.[18] More educated individuals receive less utility from leisure. Adding time *t-1* instruments maintains the strong diminishing returns on leisure (Table 1.5). In the specification without heterogeneity the negative coefficient on the interaction between health and leisure suggests they are substitutes while the positive coefficient on the interaction between health and consumption suggests they are complements. Education still lowers utility through leisure.

Table 1.6 which uses instruments from time *t*, time *t-1*, and time *t-2* produces the most precise parameter estimates. Marginal utility of consumption at the sample mean values is negative with heterogeneity but positive without heterogeneity.[19] Consumption and leisure are substitutes in both specifications. Leisure still exhibits strong diminishing returns. The specification without heterogeneity suggests consumption exhibits diminishing returns while health and consumption are substitutes. The specification with heterogeneity shows health and leisure are substitutes, and the size of the interaction between health and consumption diminishes. Consumption and leisure are decreasing in the level of education.

---

[18] The flexibility of the translog utility function has the property of producing negative marginal utility. Marginal utility of consumption with and without heterogeneity evaluated at the sample means is -0.002 and -0.004, respectively.

[19] With heterogeneity marginal utility of consumption is -0.001 and without heterogeneity marginal utility of consumption is 0.001.

Tables 1.7-1.9 replaces food consumption with nondurable consumption as the consumption measure and parallel the instrument sets used in Tables 1.4-1.6. Compared to the parameter estimates based on food consumption the alternative consumption measure tends to preserve the sign of parameter estimates, but the estimates are more precisely estimated. The trend of switching signs continues for the parameter estimates for the interaction between health and leisure. The parameter estimates suggest the effect of the interaction between health and consumption is small. As in other specifications, leisure exhibits strong diminishing returns and education lowers utility through leisure.

Tables 1.4-1.9 report test statistics for the Sargan test for the validity of overidentifying restrictions. The p-values from the Sargan test in Tables 1.4, 1.5, 1.7, and 1.8 do not reject the model specification without heterogeneity; however, they do reject the model specification with heterogeneity, suggesting invalid instruments or a misspecified utility function. The p-values in Table 1.6 and Table 1.9 based on instruments from time *t*, time *t-1*, and time *t-2* do not reject the model specification.

Tables 1.4-1.9 also report the sample mean and median values of the intertemporal substitution elasticity for consumption (ISE), a common calculation in life cycle labor supply models. The ISE is simply minus the inverse of the coefficient of relative risk aversion, or

$$ISE = \frac{U_C}{C(U_{CC})}$$

The ISE describes the proportional change in consumption expenditure needed to keep the marginal utility of wealth constant given an anticipated 1 percent change in prices.

The magnitude of the ISE increases as additional instruments are added to the instrument set. Without heterogeneity the mean and median values of the ISE range from -1.51 to -0.83 and from -1.56 to -0.88, respectively. Adding heterogeneity reduces the ISE with the exception of the specification that includes *t-2* instruments. In this specification the mean ISE is -4.08. Using the nondurable consumption measure and contemporaneous instruments with no heterogeneity gives a mean ISE of -0.91 and a median ISE of -0.92. The mean value of -0.91 implies a 1 percent equiproportionate increase in all prices leads to a 0.91 percent reduction in consumption. Adding heterogeneity produces a positive ISE, implying an increase in all prices leads to an increase in consumption. Most of the other specifications using the nondurable consumption measure and additional instrument sets imply an increase in prices lowers consumption.

Table 1.10 and Table 1.11 show income, compensated, uncompensated, and Frisch substitution elasticities based on food consumption and nondurable consumption using an instrument set containing time *t*, *t-1*, and *t-2* values (Table 1.6 and Table 1.9). The procedure to calculate the implied income, compensated, and uncompensated effects follows MaCurdy (1983). Let H represent the Hessian matrix of the utility function and $\mu = \frac{U_{c,t}}{p_t}$ be the marginal utility of income. In addition, define the price vector as $q' = (p_t, w_{i,t})$, where p_t is the price of consumption normalized to 1 and $w_{i,t}$ is the real wage. The implied income, compensated, uncompensated effects are

$$\begin{pmatrix} \frac{\partial C}{\partial Y} \\ \frac{-\partial h}{\partial Y} \end{pmatrix} = \frac{1}{n} H^{-1} q \quad \text{(Income Effects)}$$

$$\begin{pmatrix} \left.\frac{\partial C}{\partial q'}\right|_U \\ \left.\frac{-\partial h}{\partial q'}\right|_U \end{pmatrix} = \mu H^{-1} - \frac{\mu}{n} H^{-1} q q' H^{-1} \quad \text{(Compensated Effects)}$$

$$\begin{pmatrix} \left.\frac{\partial C}{\partial w}\right|_Y \\ \left.\frac{-\partial h}{\partial w}\right|_Y \end{pmatrix} = \begin{pmatrix} \left.\frac{\partial C}{\partial q'}\right|_U \\ \left.\frac{-\partial h}{\partial q'}\right|_U \end{pmatrix} + \begin{pmatrix} \frac{\partial C}{\partial Y} \\ \frac{-\partial h}{\partial Y} \end{pmatrix} h \quad \text{(Uncompensated Effects)}$$

where $n = q'H^{-1}q$. These implied effects are evaluated at the mean of the data and converted to elasticities. For these calculations full income is defined as healthy time times hourly wage; labor income is hours worked times hourly wage; and nonlabor income is full income less labor income. Since the translog utility function can produce negative marginal utilities, negative marginal utilities are set to zero. The elasticities in columns 1 and 3 are based on parameter estimates without heterogeneity while the elasticities in columns 2 and 4 are based on parameter estimates with heterogeneity. The positive nonlabor income elasticities suggest consumption is a normal good; however, for labor supply most of the nonlabor income elasticities suggest leisure is an inferior good. The labor supply elasticities near zero suggest an inelastic labor supply curve. The positive compensated and uncompensated elasticities for consumption suggest an anticipated increase in the wage leads to a rise in consumption. Most of the Frisch specific substitution elasticities are positive, implying consumption and leisure are substitutes.

1.7 Policy Simulations

A major advantage of adopting the structural approach is the ability to conduct policy simulations. Since the model connects health to wages through human capital accumulation, it can provide a sense of how much current health capital accumulation affects future wages, or how health investments today may payoff over a longer period. Specifically, it can be used to study the effects of a government subsidy on medical out-of-pocket expenditures. If health insurance reduces the cost of health care, the simplest way to represent health insurance would lower out-of-pocket medical expenditures. Figure 1.2 shows a prediction of giving individuals an additional $2,000 per year to spend on medical out-of-pocket expenditures. This additional $2,000 is considered exogenous without regard to how it is financed. The predictions are nonstochastic and are derived from the health capital and human capital production parameters. They do not account for any effects on hours worked, leisure, or consumption. The top panel of Figure 1.2 illustrates how the health stock reacts to the additional medical expenditures and provides a comparison to the actual health stock. Predicted health always exceeds actual health. The bottom panel of Figure 1.2 shows the predicted human capital stock, represented by the wage, begins at the same level as the actual human capital stock and gradually improves over time until it almost doubles compared to its initial level. These predictions suggest lowering the price of health care to spur health investment can possibly provide long run benefits in the form of increased health and human capital accumulation.

1.8 Conclusion and Future Work

This chapter presented a structural model of life cycle labor supply with a learning-by-doing technology, allowing for health to interrupt the human capital accumulation process. This model fills gaps in the health and labor literatures by introducing health to the traditional learning-by-doing model and by introducing on-the-job human capital accumulation to the health literature. Using data on male head of households from the Panel Study of Income Dynamics covering 1989-2003, the model's

Euler equations are estimated using nonlinear generalized method of moments. The estimation takes advantage of sick time data in the PSID to construct a healthy time budget constraint for each individual.

The model produces estimates of the average intertemporal substitution elasticity for consumption between -1.51 and -0.83. The results show human capital production increases with the current stock of human capital at a decreasing rate which is in accord with diminishing returns to the human capital stock. Health capital production increases with the current stock of health capital, or better current health improves future health. Health seems to have a relatively small role in interrupting on-the-job human capital accumulation. The weak role of health could be due to using self-reported health. Self-reported health measures akin to the one used in this chapter are known to be endogenous (Currie and Madrian 1999; Bound, Schoenbaum, Stinebrickner, and Waidmann 1999). Measurement error in the health will produce estimates that are biased toward zero. This measurement error in health may be mitigated with more objective health measures found in the PSID such as activities of daily living. In addition, converting the self-reported health measure to a four point scale assumes a linear relationship in health status that may not hold. For example, "excellent health" rated as 4 may not be twice "good health" rated as 2. An ordered probit can be used to relax the assumption of linearity in the health measure; however, introducing additional nonlinearity, especially in the dependent variable for health production, presents other challenges to estimating the structural model.

The model admittedly omits health insurance. The availability of health insurance is certain to influence an individual's out-of-pocket expenditures on medical care as well as access to health care. Incorporating health insurance into the model estimation can be easily done by including an indicator variable for the presence of health insurance coverage in the health production function specification. Data limitations prevent the inclusion of health insurance in the model. Unfortunately, the PSID does not always ask questions about health insurance coverage. The PSID only began asking individuals about health insurance coverage beginning in 1999 which is near the end of the sample time frame.

The health and labor supply literatures have suggested and documented differential health effects by age (Grossman 2000; Hum, Simpson, and Fissuh 2006, 2008), so future analyses should allow for heterogeneity in age. Estimating the model on different age groups is one way of introducing heterogeneity to the analysis. One possibility would estimate the model on three age groups (age 25-34; age 35-44; age 45-60).

Future work will conduct additional policy simulations, allowing for stochastic health shocks. Since the model connects health to wages through human capital accumulation, it will be used to simulate the outcomes of two hypothetical individuals who only differ in initial health status. The model can compare the wage and human capital paths over time of an individual who begins with good health to an individual who begins with poor health. Similarly, the model can compare the paths of an individual who experiences different timing of health shocks. An individual experiencing poor health early in their work career will differ from the same individual experiencing poor health later in their work career.

With so much attention on health care reform the model can provide insight into the effects of government health care policies on human capital accumulation. Specifically, the model may illustrate the human capital effects of a government subsidy on medical out-of-pocket expenditures, allowing for stochastic health shocks. If health insurance reduces the cost of health care, the simplest way to represent health insurance would lower expenditures. The model allows the study of how lowering expenditures would affect the accumulation of health capital and human capital. Lowering expenditures to spur health investment can possibly provide long run benefits in the form of increased health and human capital that has not been previously documented.

Table 1.1 Summary Statistics 1989-2003

Variable (n=15,236)	Mean	Std. Dev.	Min.	Max.
Wage (2008 dollars)	25.15	17.85	2.22	199.15
Work hours	2271.36	512.97	132	5840
Sick hours	35.04	121.92	0	2500
Leisure hours	6453.60	507.08	2920	8628
Healthy hours	8724.96	121.92	6260	8760
Health status	2.99	0.86	1	4
Food Consumption (2008 dollars)	9119.98	4132.38	557.43	48685.99
Medical Out-of-pocket (2008 dollars)	2880.47	1097.82	-441.34	20968.69
After-tax interest rate	3.94	1.25	1.70	9.30
Age	39.12	8.36	25	60
Education (years)	13.84	2.20	4	17
Number of children	1.13	1.15	0	8
Family Size	3.18	1.34	1	10
Northeast	0.20	0.40	0	1
North Central	0.31	0.46	0	1
South	0.31	0.46	0	1
West	0.18	0.38	0	1
White	0.93	0.26	0	1
Nonwhite	0.07	0.26	0	1
Married	0.83	0.38	0	1

Table 1.1 (continued)

Variable	Sick Hours Equal to Zero	Sick Hours Between Zero and 75th Percentile	Sick Hours Greater than 75th Percentile
Wage (2008 dollars)	25.07	25.84	25.19
Work hours	2301.04	2305.26	2208.75
Sick hours	0.00	16.02	105.58
Leisure hours	6458.96	6438.72	6445.68
Healthy hours	8760.00	8743.98	8654.42
Health status	3.03	3.14	2.90
Food Consumption (2008 dollars)	9237.68	9040.64	8905.94
Medical Out-of-pocket (2008 dollars)	2891.48	2947.29	2848.55
After-tax interest rate	3.88	3.94	4.05
Age	39.66	37.56	38.31
Education (years)	13.71	14.28	14.02
Number of children	1.12	1.27	1.13
Family Size	3.19	3.28	3.13
Northeast	0.19	0.20	0.22
North Central	0.32	0.35	0.31
South	0.32	0.29	0.28
West	0.17	0.16	0.19
White	0.92	0.96	0.93
Nonwhite	0.08	0.04	0.07
Married	0.84	0.87	0.80

Table 1.2 Human Capital Production Parameters

Parameter	(1) Estimate	(2) Std. Error	(3) Estimate	(4) Std. Error
α_0 (w_t)	0.976***	(0.108)	1.174***	(0.205)
α_0 (w_t) (educ)			-0.00313	(0.0119)
α_0 (w_t) (nonwhite)			-0.0537	(0.0702)
α_1 (w_t^2)	-0.393***	(0.0393)	-1.200***	(0.259)
α_1 (w_t^2) (educ)			0.0586***	(0.0174)
α_1 (w_t^2)(nonwhite)			-0.0308	(0.103)
α_2 (w_tN_t)	0.0395	(0.0366)	0.00219	(0.0354)
α_3 (N_t)	-0.00486	(0.0119)	0.0371	(0.0370)
α_3 (N_t)(educ)			-0.00222	(0.00272)
α_3 (N_t)(nonwhite)			0.0189	(0.0203)
α_4 (N_t^2)	0.00229	(0.00225)	-0.00824	(0.00788)
α_4 (N_t^2)(educ)			0.000674	(0.000633)
α_4 (N_t^2)(nonwhite)			-0.00421	(0.00419)
α_5 (H_t)	0.00224	(0.00812)	-0.0103	(0.0343)
α_5 (H_t)(educ)			0.00131	(0.00249)
α_5 (H_t) (nonwhite)			-0.00591	(0.0191)
α_6 (H_t^2)	-0.00148	(0.00132)	0.000847	(0.00644)
α_6 (H_t^2)(educ)			-0.000124	(0.000482)
α_6 (H_t^2)(nonwhite)			0.000964	(0.00328)
α_7 (H_tN_t)	-0.00146	(0.00310)	-0.000776	(0.00316)
α_8 (w_tH_t)	0.0553**	(0.0241)	0.00410	(0.0257)
τ_{1989}	-0.00318	(0.0219)	-0.0331	(0.0263)
τ_{1990}	0.00635	(0.0222)	-0.0239	(0.0263)
τ_{1991}	0.00266	(0.0223)	-0.0266	(0.0267)
τ_{1992}	0.0124	(0.0220)	-0.0178	(0.0264)
τ_{1993}	-0.00140	(0.0218)	-0.0315	(0.0263)
τ_{1994}	-0.00234	(0.0218)	-0.0333	(0.0262)
τ_{1995}	-0.00317	(0.0218)	-0.0343	(0.0262)
τ_{1996}	0.00381	(0.0219)	-0.0273	(0.0263)
τ_{1997}	0.0176	(0.0219)	-0.0133	(0.0262)
τ_{1998}	0.0174	(0.0219)	-0.0131	(0.0264)
τ_{1999}	0.00505	(0.0219)	-0.0252	(0.0263)
Observations	15236		15236	
R-squared	0.884		0.892	

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1
Columns 3 and 4 include observed heterogeneity.

Table 1.3 Health Capital Production Parameters

Parameter	(1) Estimate	(2) Std. Error	(3) Estimate	(4) Std. Error
θ_0 (H_t)	0.450***	(0.0944)	0.562**	(0.241)
θ_0 (H_t) (educ)			-0.00690	(0.0162)
θ_0 (H_t) (nonwhite)			0.240*	(0.132)
θ_1 (H_t^2)	-0.00904	(0.00673)	-0.0436	(0.0406)
θ_1 (H_t^2)(educ)			0.00266	(0.00292)
θ_1 (H_t^2)(nonwhite)			-0.0445*	(0.0249)
θ_2 (H_tL_t)	0.0315**	(0.0126)	0.0274**	(0.0129)
θ_3 (L_t)	-0.167	(0.126)	-0.226	(0.195)
θ_3 (L_t)(educ)			0.00835	(0.00950)
θ_3 (L_t)(nonwhite)			-0.0593	(0.0698)
θ_4 (L_t^2)	-0.000181	(0.00903)	0.00293	(0.0180)
θ_4 (L_t^2)(educ)			-0.000384	(0.00110)
θ_4 (L_t^2)(nonwhite)			0.00316	(0.00759)
θ_5 (M_t)	-1.098*	(0.652)	-0.461	(1.075)
θ_5 (M_t)(educ)			-0.0444	(0.0509)
θ_5 (M_t)(nonwhite)			-1.276*	(0.685)
θ_6 (M_t^2)	-0.126	(0.0994)	-0.898	(1.209)
θ_6 (M_t^2)(educ)			0.0627	(0.0732)
θ_6 (M_t^2)(nonwhite)			2.618*	(1.451)
θ_7 (H_tM_t)	0.173***	(0.0622)	0.0911	(0.0731)
θ_8 (L_tM_t)	0.149	(0.0950)	0.111	(0.111)
τ_{1989}	2.099***	(0.495)	1.965***	(0.522)
τ_{1990}	2.084***	(0.495)	1.956***	(0.522)
τ_{1991}	2.091***	(0.494)	1.960***	(0.522)
τ_{1992}	2.099***	(0.494)	1.958***	(0.522)
τ_{1993}	2.050***	(0.494)	1.927***	(0.522)
τ_{1994}	2.078***	(0.495)	1.947***	(0.522)
τ_{1995}	2.065***	(0.495)	1.936***	(0.522)
τ_{1996}	2.107***	(0.494)	1.971***	(0.522)
τ_{1997}	2.104***	(0.494)	1.953***	(0.521)
τ_{1998}	2.036***	(0.494)	1.884***	(0.522)
τ_{1999}	2.093***	(0.494)	1.948***	(0.522)
Observations	15236		15236	
R-squared	0.954		0.955	

Robust standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$
Columns 3 and 4 include observed heterogeneity.

Table 1.4 Utility Parameters Using Time t Instruments

Parameter	(1) Estimate	(2) Std. Error	(3) Estimate	(4) Std. Error
γ_0 (ln L_t)	1.00		1.00	
γ_0 (ln L_t) (educ)			-0.0097***	(0.0023)
γ_0 (ln L_t) (nonwhite)			-0.1668	(0.4874)
γ_1 (ln C_t)	0.0016	(0.0013)	-0.0015***	(0.0007)
γ_1 (ln C_t) (educ)			-0.00001	(0.00001)
γ_1 (ln C_t) (nonwhite)			-0.0117	(0.0166)
γ_2 (ln L_t) (ln C_t)	-0.0009	(0.0007)	0.0009***	(0.0003)
γ_3 (ln L_t)^2	-0.2732***	(0.0084)	-0.2326***	(0.0083)
γ_4 (ln C_t)^2	0.00001	(0.00002)	0.0076	(0.0781)
γ_6 (ln H_t) (ln L_t)	0.0238	(0.0225)	-0.0062	(0.0126)
γ_7 (ln H_t) (ln C_t)	-0.0002	(0.0002)	0.0001	(0.00003)
ISE (mean)	-0.83		-0.02	
ISE (median)	-0.88		-0.15	
Sargan test [df]	36.34[31]		43.45[27]	
(p-value)	(0.23)		(0.02)	
Observations	15236		15236	

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1
Columns 3 and 4 include observed heterogeneity.

Table 1.5 Utility Parameters Using Time t and t-1 Instruments

Parameter	(1) Estimate	(2) Std. Error	(3) Estimate	(4) Std. Error
γ_0 (ln L_t)	1.00		1.00	
γ_0 (ln L_t) (educ)			-0.0049***	(0.0008)
γ_0 (ln L_t) (nonwhite)			-0.1250	(0.0821)
γ_1 (ln C_t)	0.0006	(0.0005)	-0.0001	(0.0004)
γ_1 (ln C_t) (educ)			0	(0)
γ_1 (ln C_t) (nonwhite)			-0.0018	(0.0022)
γ_2 (ln L_t) (ln C_t)	-0.0003	(0.0003)	0.0001	(0.0002)
γ_3 (ln L_t)^2	-0.2602***	(0.0036)	-0.2517***	(0.0028)
γ_4 (ln C_t)^2	0.00001	(0.00001)	0.0005	(0.0034)
γ_6 (ln H_t) (ln L_t)	-0.0368***	(0.0121)	0.0103	(0.0083)
γ_7 (ln H_t) (ln C_t)	0.0002***	(0.0001)	-0.00001	(0.00001)
ISE (mean)	-1.08		0.01	
ISE (median)	-1.10		-0.16	
Sargan test [df]	56.92[65]		77.59[61]	
(p-value)	(0.75)		(0.07)	
Observations	12903		12903	

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1
Columns 3 and 4 include observed heterogeneity.

Table 1.6 Utility Parameters Using Time t, t-1, and t-2 Instruments

Parameter	(1) Estimate	(2) Std. Error	(3) Estimate	(4) Std. Error
γ_0 (ln L_t)	1.00		1.00	
γ_0 (ln L_t) (educ)			-0.0023***	(0.0009)
γ_0 (ln L_t) (nonwhite)			-0.1264	(0.2984)
γ_1 (ln C_t)	0.0049***	(0.0009)	0.0051***	(0.0005)
γ_1 (ln C_t) (educ)			-0.0001***	(0.00001)
γ_1 (ln C_t) (nonwhite)			0.0133	(0.0191)
γ_2 (ln L_t) (ln C_t)	-0.0028***	(0.0005)	-0.0023***	(0.0002)
γ_3 (ln L_t)^2	-0.2735***	(0.0032)	-0.2385***	(0.0054)
γ_4 (ln C_t)^2	-0.0002***	(0.00003)	-0.0003	(0.0004)
γ_6 (ln H_t) (ln L_t)	-0.0018	(0.0104)	-0.0589***	(0.0149)
γ_7 (ln H_t) (ln C_t)	-0.0006***	(0.0002)	0	(0.00002)
ISE (mean)	-1.51		-4.08	
ISE (median)	-1.56		-0.03	
Sargan test [df]	97.37[96]		87.73[92]	
(p-value)	(0.44)		(0.61)	
Observations	10570		10570	

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1
Columns 3 and 4 include observed heterogeneity.

Table 1.7 Utility Parameters Using Time t Instruments and Nondurable Consumption

Parameter	(1) Estimate	(2) Std. Error	(3) Estimate	(4) Std. Error
γ_0 (ln L_t)	1.00		1.00	
γ_0 (ln L_t) (educ)			-0.0084***	(0.0023)
γ_0 (ln L_t) (nonwhite)			-0.1474	(0.3622)
γ_1 (ln C_t)	0.0038**	(0.0021)	-0.0024***	(0.0012)
γ_1 (ln C_t) (educ)			0.00001	(0.00001)
γ_1 (ln C_t) (nonwhite)			-0.0268	(0.0264)
γ_2 (ln L_t) (ln C_t)	-0.0020**	(0.0011)	0.0014***	(0.0006)
γ_3 (ln L_t)^2	-0.2781***	(0.0079)	-0.2356***	(0.0080)
γ_4 (ln C_t)^2	0.00002	(0.00003)	0.0795	(0.1120)
γ_6 (ln H_t) (ln L_t)	0.0366**	(0.0214)	-0.0126	(0.0119)
γ_7 (ln H_t) (ln C_t)	-0.0005**	(0.0003)	0.0001**	(0.0001)
ISE (mean)	-0.91		0.76	
ISE (median)	-0.92		0.30	
Sargan test [df]	34.98 [31]		41.89 [27]	
(p-value)	(0.28)		(0.03)	
Observations	15236		15236	

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1
Columns 3 and 4 include observed heterogeneity.

Table 1.8 Utility Parameters Using Time t and t-1 Instruments and Nondurable Consumption

Parameter	(1) Estimate	(2) Std. Error	(3) Estimate	(4) Std. Error
γ_0 (ln L_t)	1.00		1.00	
γ_0 (ln L_t) (educ)			-0.0051***	(0.0008)
γ_0 (ln L_t) (nonwhite)			-0.2011	(0.1264)
γ_1 (ln C_t)	0.0006	(0.0009)	-0.0004	(0.0007)
γ_1 (ln C_t) (educ)			-0.00001	(0.00001)
γ_1 (ln C_t) (nonwhite)			-0.0043	(0.0067)
γ_2 (ln L_t) (ln C_t)	-0.0003	(0.0005)	0.0003	(0.0003)
γ_3 (ln L_t)^2	-0.2609***	(0.0038)	-0.2502***	(0.0026)
γ_4 (ln C_t)^2	0.00001	(0.00001)	-0.0032	(0.0048)
γ_6 (ln H_t) (ln L_t)	-0.0342***	(0.0130)	0.0080	(0.0081)
γ_7 (ln H_t) (ln C_t)	0.0003***	(0.0001)	0	(0.00001)
ISE (mean)	-1.14		-3.66	
ISE (median)	-1.08		0.39	
Sargan test [df]	58.65[65]		76.85[61]	
(p-value)	(0.70)		(0.08)	
Observations	12903		12903	

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1
Columns 3 and 4 include observed heterogeneity.

Table 1.9 Utility Parameters Using Time t, t-1, and t-2 Instruments and Nondurable Consumption

Parameter	(1) Estimate	(2) Std. Error	(3) Estimate	(4) Std. Error
γ_0 (ln L_t)	1.00		1.00	
γ_0 (ln L_t) (educ)			-0.0033***	(0.0009)
γ_0 (ln L_t) (nonwhite)			0.0252	(0.4233)
γ_1 (ln C_t)	0.0123***	(0.0021)	0.0097	(0.0011)
γ_1 (ln C_t) (educ)			-0.0001***	(0.00002)
γ_1 (ln C_t) (nonwhite)			0.0051	(0.0443)
γ_2 (ln L_t) (ln C_t)	-0.0069***	(0.0013)	-0.0044***	(0.0004)
γ_3 (ln L_t)^2	-0.2790***	(0.0039)	-0.2369***	(0.0055)
γ_4 (ln C_t)^2	-0.0005***	(0.0001)	-0.00003	(0.0008)
γ_6 (ln H_t) (ln L_t)	0.0164	(0.0125)	-0.0473***	(0.0153)
γ_7 (ln H_t) (ln C_t)	-0.0010***	(0.0003)	-0.00002	(0.00004)
ISE (mean)	-5.70		-1.30	
ISE (median)	-1.75		-1.12	
Sargan test [df]	98.32[96]		96.10[92]	
(p-value)	(0.42)		(0.36)	
Observations	10570		10570	

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1
Columns 3 and 4 include observed heterogeneity.

Table 1.10 Elasticities Based On Food Consumption

	Consumption		Labor	
Elasticity	(1)	(2)	(3)	(4)
Income Elasticity	1.52	1.83	0.01	-0.02
Compensated Elasticity	0.00002	0.006	0.001	0.02
Uncompensated Elasticity	1.13	0.73	0.001	0.02
Frisch specific substitution elasticity	0.01	-0.03	0.001	0.02

Elasticities based on parameters from model with *t*, *t-1*, and *t-2* instruments and are evaluated at the mean values of the data. Columns 1 and 3 do not include heterogeneity. Columns 2 and 4 include observed heterogeneity.

Table 1.11 Elasticities Based On Nondurable Consumption

	Consumption		Labor	
Elasticity	(1)	(2)	(3)	(4)
Income Elasticity	0.95	0.88	0.02	0.01
Compensated Elasticity	0.00002	0.0007	0.002	0.01
Uncompensated Elasticity	0.72	0.40	0.002	0.01
Frisch specific substitution elasticity	0.04	0.01	0.003	0.01

Elasticities based on parameters from model with *t*, *t-1*, and *t-2* instruments and are evaluated at the mean values of the data. Columns 1 and 3 do not include heterogeneity. Columns 2 and 4 include observed heterogeneity.

Figure 1.1 Distribution of Sick Hours

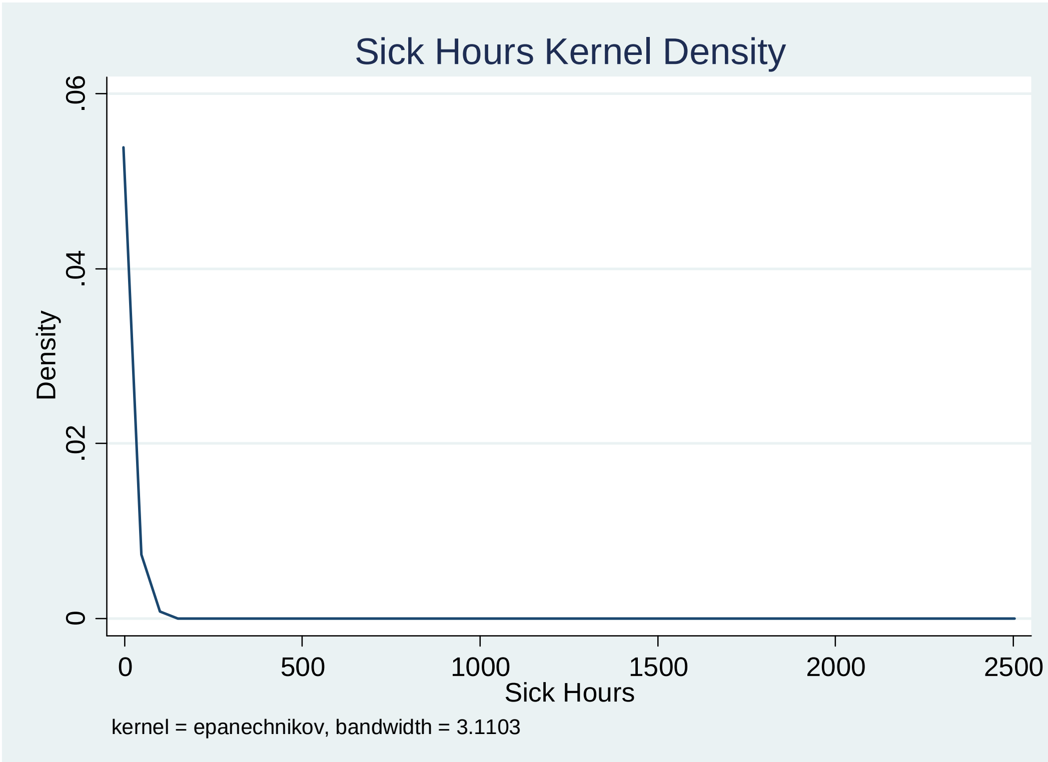
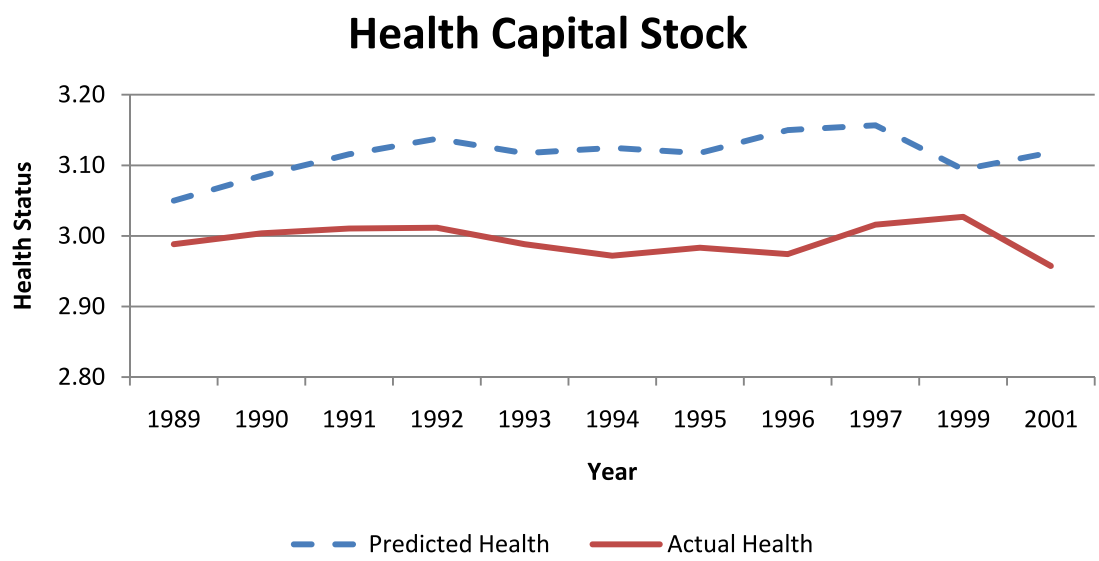
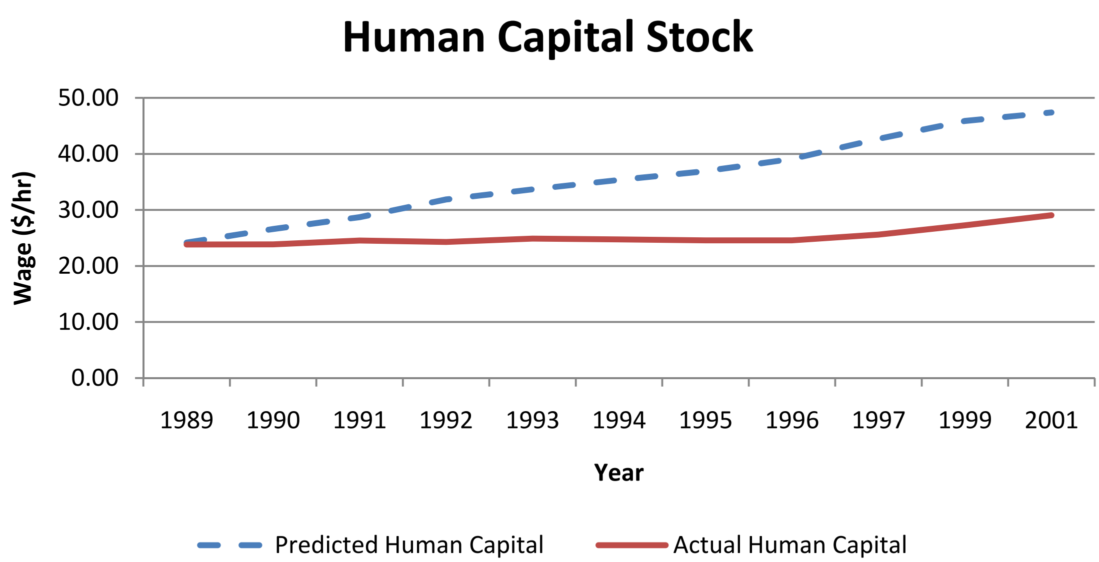


Figure 1.2 Prediction of Additional Medical Out-of-Pocket Expenditures





2 NONCOGNITIVE SKILLS AND THE RACIAL WAGE GAP

2.1 Introduction

Economists have established the importance of cognitive ability and human capital in determining the returns to education and other behaviors. Sociologists and psychologists have focused on the role of noncognitive skills in social outcomes. Noncognitive skills refers to a type of human capital, or "psychological capital," describing a person's self-perception, work ethic, ethical orientation, and overall outlook on life.[20] Common sense suggests these noncognitive skills certainly influence an individual's productivity along with cognitive skills. Economists typically account for noncognitive skills in an error term of an estimating equation, claiming personality traits are difficult to measure or are just unobservable. They address these unobserved skills using an error component model on panel data that relies on fixed effects or random effects. Sociologists and psychologists have constructed measures of noncognitive skills that allow a researcher to control for some of the unobserved heterogeneity.

Much effort has been expended on studying the racial wage gap. Common explanations for the racial wage gap include employer and consumer discrimination, varying school quality, and differences in premarket factors. Neal and Johnson (1996) show the black-white wage gap shrinks after including a premarket factor, cognitive skills, in a parsimonious wage equation. While it is apparent that cognitive skills are an important premarket factor to consider, it seems natural that noncognitive skills may be an important factor as well. Economists are developing a better understanding of the importance of these skills for educational attainment and economic success in the general

[20] ter Weel (2008) and Heckman (2007) expand the definition of noncognitive skills beyond psychological and behavioral traits to include time preference, risk aversion, and preference for leisure.

population (Coleman and DeLeire, 2003; Heckman, Stixrud, and Urzua 2006). Less understood is the impact of these skills among subgroups of the general population, specifically racial groups.

This chapter investigates the role of noncognitive skills in explaining racial gaps in wages. Noncognitive skills, measured by locus of control and self-esteem, are added to a simple wage regression from Neal and Johnson (1996) to examine their effect on the wage gap. The analysis extends the racial wage gap and noncognitive skills literatures by studying the effect of noncognitive skills on wage gaps across the entire wage distribution. Using data from the National Longitudinal Survey of Youth 1979 (NLSY79) spanning 1991-2006 this chapter estimates wage regressions based on three estimators: a pooled estimator, a between estimator, and a quantile estimator. The wage regressions take advantage of the timing of when noncognitive skills and wages are measured in the NLSY79. The wage regressions relate cognitive and noncognitive skills measured at the beginning of the NLSY79 before individuals enter the labor market or begin post secondary schooling to wages measured later in life. The various model specifications capture the separate and simultaneous effects of cognitive and noncognitive skills on the wage gap.

Ordinary least squares results show noncognitive skills account for a small portion of the male black-white wage gap when measured at the mean of the wage distribution. Noncognitive skills have differing effects for black women. Ordinary least squares results show locus of control shrinks the gap, but self-esteem widens it. After controlling for cognitive and noncognitive skills, there still exists a significant return to noncognitive skills. External individuals earn less, and individuals with higher self-esteem earn more. Quantile regressions controlling for cognitive and noncognitive skills suggest the black-white male wage gap persists at all points of the wage distribution while the black-white female wage gap exists at only the highest portion of the wage distribution. Hispanic men earn less than white men at lower quantiles but earn more at higher quantiles. After controlling for cognitive and noncognitive skills, Hispanic women always earn more than white women across the entire wage distribution. In addition, after controlling for cognitive skills and noncognitive skills, the return to cognitive skills exceeds the return to noncognitive skills across the entire wage distribution.

2.2 Noncognitive Literature and Racial Wage Gaps

Economists began studying the role of noncognitive skills over three decades ago. Bowles and Gintis (1976) find that low skill markets contain employers that place a higher value on noncognitive skills. Edwards (1976) finds blue-collar supervisors prefer these skills over cognitive skills, while Mueser (1979) shows noncognitive skills are just as important as cognitive skills in determining wages.

Andrisani (1977) specifically examines the effect of locus of control on wages and occupational attainment in black and white men. Locus of control, measured by the Rotter Scale, gauges the degree of internal or external control an individual has over their life. It describes the extent to which individuals believe they have control over their lives through self-motivation or self-determination (internal control) as opposed to the extent that the environment (chance, fate, luck) controls their lives (external control) (National Longitudinal Survey of Youth 2007). Andrisani uses a sample of young and middle aged men taken from the National Longitudinal Survey in 1968 and 1969. He studies how locus of control relates to wages and occupation two years after they are measured to determine the subsequent effect of attitude. He includes locus of control in a standard earnings equation with common controls (education, tenure, experience, region, urban/rural, etc.). Locus of control among both racial groups has similar payoffs--more internal individuals have higher wages. Differences exist in occupational advancement. Younger white men experience a stronger effect of being internal on occupational advancement than younger black men.

Duncan and Morgan (1981) replicate Andrisani's work with data from the Panel Study of Income Dynamics (PSID). Since the PSID does not directly ask the Rotter Scale, Duncan and Morgan match answers to open ended questions coded by PSID staff in 1968 to components of the Rotter Scale and use the answers as a measure of self-efficacy. They examine the effects of self-efficacy on wages two and four years later. Unlike Andrisani, they find no significant effect on either young black or young white men two years later; however, they find a positive significant effect for white men four years later.

Goldsmith, Veum, and Darity (1997) recognize that wages and self-esteem are determined jointly, so they estimate two equations with locus of control and self-esteem data from the 1987 and 1980 NLSY79. In their specification, self-esteem directly enters the wage equation while locus of control directly enters the self-esteem equation, so locus of control only indirectly affects wages through self-esteem. Their equations are identified through strong exclusion restrictions. The wage equation leaves out locus of control, and the self-esteem equation leaves out local labor market conditions. Unlike previous studies, Goldsmith, Veum, and Darity control for cognitive ability through the

1980 Armed Forces Qualifying Test (AFQT) score. Their findings suggest self-esteem has a stronger positive effect on wages than human capital, and locus of control significantly affects self-esteem. They interpret their findings as "psychological capital" affecting wages in two ways: directly through self-esteem and indirectly through locus of control. In addition, they find blacks earn less, and higher wages lead to better self-esteem.

Coleman and DeLeire (2003) develop a theoretical model connecting locus of control among teenagers to educational attainment through expectations on the return to education. Their model implies more internal teenagers, who believe their current actions influence future outcomes, are more likely to make investments in education. Their model offers a test of whether locus of control is just a proxy for ability. If locus of control simply proxies for ability, then internal teenagers, both high school dropouts and high school graduates, will expect better outcomes. Coleman and DeLeire test their theory with the National Educational Longitudinal Study (NELS) data by regressing educational attainment (high school or college completion) on locus of control measured in the eighth grade. While they show internal teenagers are more likely to graduate from high school, they do not explain any racial differences in this effect. Race variables only enter as additional controls in their estimation.

In a replication study with different data Cebi (2007) tests Coleman and DeLeire's model with educational attainment data from the NLSY79 and examines the effects of noncognitive skills on wages. Cebi uses locus of control measured in 1979 to explain the probability of graduating from high school and attending college in 1982. Her results differ from Coleman and DeLeire. After controlling for AFQT score, she finds no evidence that locus of control predicts high school graduation or attending college. She also finds a small significant return to locus of control in year 2000 wages, so more internal individuals earn more later in life. Also, the black wage gap shrinks after accounting for AFQT and locus of control. Cebi's analysis focuses on a pooled sample of men and women and studies the effect of noncognitive skills at the mean of the wage distribution. Cebi only considers locus of control as the measure of noncognitive skills.

Carneiro, Heckman, and Masterov (2005) document differences in noncognitive skills among black, Hispanic, and white children as measured by the antisocial behavior index in the children of the NLSY79 cohort. They show environmental differences account for the majority of the minority-white gap in noncognitive skills. This work only addresses the differences in early childhood and does not relate these differences to wages.

Heckman, Stixrud, and Urzua (2006) provide an extensive treatment of noncognitive skills. They develop a statistical model to describe the importance of cognitive and noncognitive skills in determining schooling, work experience, wages, occupational choice, and a number of risky behavioral outcomes. Though they do not address differences among racial groups, they advance the economic noncognitive literature in two important ways. First, they consider the simultaneous effects of cognitive and noncognitive skills on a variety of outcomes beyond just the standard labor market and educational outcomes. Second, they develop a methodology that accounts for the endogeneity of schooling and measurement error in test scores. In this context schooling may cause higher test scores. Their methodology uses a common set of latent cognitive and noncognitive factors to determine each outcome of interest. In addition, they estimate a test score equation for each cognitive and noncognitive measure that depends on the level of schooling at the time of the test and the appropriate latent factor. Allowing schooling to enter this equation controls for its influence on the test score. They utilize the AFQT score, Rotter Scale, and Rosenberg Self-Esteem score from the NLSY79 in estimating their model. They present evidence that schooling affects both measures of cognitive and noncognitive abilities, so it is important to control for this effect. They focus their analysis on the differences between men and women, not racial differences.

Urzua (2008) estimates a structural model of schooling choice, labor market behavior, and incarceration to examine the importance of unobserved cognitive and noncognitive abilities in explaining the black-white gaps in these outcomes. Specific labor market behavior includes wages, earnings, and hours worked. The model addresses the endogeneity of schooling choice because individuals make schooling decisions based on differences in returns to schooling. Measurement error in cognitive and noncognitive abilities is handled in a similar way as Heckman, Stixrud, and Urzua (2006), and the analysis uses the same measures of these skills in the NLSY79 as in Heckman, Stixrud, and Urzua (2006). Urzua finds black-white differences in the unobserved abilities for both cognitive and noncognitive distributions. For schooling choices, hours worked, and wages, noncognitive abilities matter more for blacks than for whites. Unobserved noncognitive abilities do not account for much of the black-white wage or earnings gap; however, unobserved noncognitive abilities do play a stronger role in explaining the gap in incarceration rates. Urzua simulates the effect of assigning blacks white characteristics to study how the gaps in wages, earnings, schooling, and incarceration change. When blacks have the white distribution of unobserved cognitive abilities, they attain equal or better education levels as whites, and the gap in wages and earnings falls by about 40 percent, smaller than the literature which claims a 50-75 percent reduction when observed cognitive ability is controlled for. Urzua only studies the black-white wage and earnings gap at the mean and does not consider the gap at other points of the wage or earnings distributions. Giving blacks the white distribution of unobserved noncognitive abilities does not change the wage or earnings gap by much. So, it is unobserved cognitive ability that explains racial gaps in schooling attainment and labor market outcomes. When blacks have the white distribution of unobserved cognitive and noncognitive abilities, they achieve the lowest level of incarceration rates.

This chapter presents wage regressions relating noncognitive skills measured during the teenage years before individuals enter the labor market or begin post secondary schooling to wages measured later in life. This chapter extends the noncognitive literature in several ways. First, the analysis examines the effect of noncognitive skills on the wage gap for Hispanics, not just blacks. Second, the analysis considers the racial wage gap for each gender and race combination. The noncognitive literature has typically focused on differences between men and women without examining racial differences within gender. Third, the analysis extends the racial wage gap and noncognitive skills literatures by going beyond the wage gap measured at the mean of the wage distribution. The analysis studies the effect of noncognitive skills on wage gaps across the entire wage distribution.

2.3 Model Specification

The model specification for this chapter draws from the literature on race differences in premarket human capital and wages (O'Neill 1990; Maxwell 1994; Carneiro, Heckman, and Masterov 2005). Neal and Johnson (1996) carefully test a theory in this literature that relates the black-white and Hispanic-white wage gap to differences in the skills measured by the Armed Forces Qualifying Test (AFQT) at labor market entry. The specification for this chapter incorporates noncognitive skills as an additional premarket factor in the model presented by Neal and Johnson (1996). The simple specification is of the following form:

$$\ln wage_{i,t} = \beta_0 + \beta_1 Black_i + \beta_2 Hispanic_i + \beta_3 Age_{i,t} + \beta_4 AFQT_{i,1980} + \beta_5 AFQT^2_{i,1980} + \beta_6 Noncog_{i,1979/1980} + \beta_7 Noncog^2_{i,1979/1980} + \varepsilon_i$$

where $wage_{i,t}$ is the real wage of person *i* in year *t* in 2009 dollars, adjusted by the personal consumption expenditures price index. $Black_i$ and $Hispanic_i$ are dummy variables for black and Hispanic racial groups (white is the omitted category) while Age_i is the person's age. $AFQT_{i,1980}$ is the score from AFQT in 1980 and serves as a measure of cognitive skills. The AFQT is constructed from summing scores on sections 2-5 of the Armed Services Vocational Aptitude Battery Test (ASVAB): arithmetic reasoning, word knowledge, paragraph comprehension, and numerical operations. The raw AFQT score is then normalized to have a mean of zero and standard deviation of one.

$Noncog_i$ is a noncognitive measure in 1979 or 1980. The AFQT score is commonly used by economists, but the measures of noncognitive skills are less common and warrant further discussion. Two measures of noncognitive skills are used: the Rotter Scale and the Rosenberg Self-Esteem Scale. The Rotter Internal-External Locus of

Control Scale, administered in 1979, is a four item questionnaire designed to measure the degree to which a person has control over their life through self-motivation or self-determination (internal control) as opposed to the extent that the environment (i.e., chance, fate, luck) controls their life (external control) (NLSY documentation 2007). A higher score reflects a more external person. The four item questionnaire consists of these statement pairs listed below:

1. Rotter 1
   a. What happens to me is my own doing
   b. Sometimes I feel that I don't have enough control over the direction my life is taking
2. Rotter 2
   a. When I make plans, I am almost certain that I can make them work
   b. It is not always wise to plan too far ahead, because many things turn out to be a matter of good or bad fortune anyhow
3. Rotter 3
   a. Getting what I want has little or nothing to do with luck
   b. Many times we might just as well decide what to do by flipping a coin
4. Rotter 4
   a. Many times I feel that I have little influence over the things that happen to me
   b. It is impossible for me to believe that chance or luck plays an important role in my life

The first statement in each pair corresponds to an internal control item while the second statement corresponds to an external control item. A person chooses one of the paired statements and decides if the chosen statement is much closer or slightly closer to their opinion of themselves. Together these two answers generate a four point scale for each paired item. The Rotter score is the average over the four paired items (Rotter 1, Rotter 2, Rotter 3, and Rotter 4). The Rotter score is normalized in the same way as the AFQT score.[21]

The Rosenberg Self-Esteem Scale, administered in 1980, is a 10 item scale that measures the self-evaluation that an individual makes and characterizes the degree of approval or disapproval toward oneself (NLSY documentation 2007). A higher score corresponds to higher self-esteem. A person answers the following ten statements of approval or disapproval with strongly agree, agree, disagree, or strongly disagree:

1. I feel I'm a person of worth, at least on an equal basis with others
2. I feel that I have a number of good qualities
3. All in all, I am inclined to feel that I am a failure
4. I am able to do things as well as most other people
5. I feel I do not have much to be proud of
6. I take a positive attitude toward myself
7. On the whole, I am satisfied with myself
8. I wish I could have more respect for myself
9. I certainly feel useless at times

[21] This averaging of the paired item scores and normalization of the Rotter score follows Heckman, Stixrud, and Urzua (2006).

10. At times I think I am no good at all

The Rosenberg score averages the responses over the ten statements. This average is also normalized to have mean zero and standard deviation one.

Four variations of the specification are estimated:

1. Black, Hispanic, Age
2. Black, Hispanic, Age, AFQT, $AFQT^2$
3. Black, Hispanic, Age, Noncog, $Noncog^2$
4. Black, Hispanic, Age, AFQT, $AFQT^2$, Noncog, $Noncog^2$

These specifications are meant to capture the separate and simultaneous effects of cognitive and noncognitive abilities. As a robustness check, specifications that replace AFQT with components of AFQT are estimated. Specifications that interact race with cognitive skills and noncognitive skills to allow for differing returns by race are estimated. In addition, specifications that control for region (South and nonsouth) and interact region with cognitive and noncognitive skills are estimated.

The key idea from Neal and Johnson (1996) is these factors are measured *before* labor market entry to eliminate any effects due to worker choices or labor market discrimination. These specifications omit education, experience, and occupation, commonly included regressors in an earnings specification, because they are also endogenous. Including these common regressors biases the effect of race on wages if discrimination against blacks or Hispanics causes them to make occupation or education choices different from whites. Following Neal and Johnson the sample is limited to individuals born after 1961 who are 18 or younger when they took the ASVAB test. Individuals in this group most likely have not entered the labor market or begun postsecondary schooling when they took the ASVAB test. Not including individuals who are over 18 eliminates any influence of schooling or the labor market on the AFQT score. In addition, following Neal and Johnson (1996) only individuals with wage observations between $1 and $75 are considered.

2.4 Estimation Methods

The variations of the simple model specification are estimated using a pooled estimator, a between estimator, and a quantile estimator. The pooled estimator includes annual time dummy variables for 1992-2006, so the specification becomes

$$\ln wage_{i,t} = \boldsymbol{x}'_{i,t}\beta + \varepsilon_{i,t}$$

This specification is estimated using ordinary least squares with clustered standard errors to correct for the longitudinal structure of the NLSY. The standard errors account for repeated observations of individuals over time.

The between estimator is the ordinary least squares estimator on individual time means of the data, or

$$\overline{\ln wage_{\iota}} = \overline{\boldsymbol{x}}'_{\boldsymbol{i}}\beta + \bar{\varepsilon}_i$$

where $\ln wage_i = \frac{1}{T}\sum_{t=1}^{T} \ln wage_{i,t}$ and similarly for other variables. The between estimator averages the individual data over time, keeping one observation per individual. This smoothing offers the advantage of reducing any measurement error associated with the wage and an improvement in efficiency. This specification is also estimated using ordinary least squares but with heteroskedastic robust standard errors.

Unlike the ordinary least squares estimator which estimates the conditional mean, the quantile estimator estimates the conditional quantile of the wage as a linear function of the observables. Formally, the quantile estimator solves the minimization problem

$$\min_{\beta} \sum \rho_\tau(\ln wage_{i,t} - \boldsymbol{x}_{\boldsymbol{i,t}}'\beta)$$

where $\rho_\tau(\cdot)$ represents the τ'th quantile "check" function, or absolute value function. Each specification is estimated for deciles $\tau = \{.10, .20, .30, .40, .50, .60, .70, .80, .90\}$ on the pooled data and the time-averaged data used for the between estimator. Standard errors are estimated using the nonparametric bootstrap with 100 replications.

2.5 Data

The data in the analysis come from the National Longitudinal Survey of Youth 1979 (NLSY79). The NLSY79 contains 12,686 individuals between the ages of 14 and 21 at the time of the first interview in 1979. The NLSY79 collects information on labor market outcomes as well as cognitive and noncognitive abilities. The interviews occur every year for 1979-1994 and every two years for 1996-2006. This analysis uses wage observations beginning with 1991 when sample ages were 26-29 and ending with 2006 when sample ages were 41-45. The NLSY79 reports an hourly wage if the individual is an hourly worker and reports an hourly wage; otherwise, the NLSY79 calculates an hourly wage rate from earnings and hours worked (NLSY documentation 2010). As mentioned before, the cognitive measure comes from the AFQT score calculated from the ASVAB taken in 1980. The noncognitive measures come from the Rotter Scale for locus of control and Rosenberg Self-Esteem Scale administered in 1979 and 1980, respectively. Each specification is estimated for men and women separately and covers the years 1991-1994 (annually) and 1996-2006 (biennially).

Table 2.1 provides summary statistics for the entire analysis sample, and Table 2.2 shows summary statistics by gender and race. Relatively more men than women comprise the sample with the majority of the sample white (52%) followed by black (29%) and Hispanic (19%). Blacks earn an average hourly wage below the sample average hourly wage of $16.72 (2009 dollars) while whites earn an average wage above it. Figure 2.1 displays kernel density estimates of the standardized AFQT, Rotter, and Rosenberg scores by racial group and by gender. The horizontal and vertical scales are the same for ease of comparison. The white AFQT distribution is clearly shifted to the right when compared to the other distributions. On average whites scored highest on the AFQT test (standardized average .32), and blacks scored the lowest (standardized average -.45). The black and Hispanic Rotter distributions share a similar shape while the white Rotter distribution contains more mass less than zero, suggesting whites are more internal than the other groups. The Rosenberg Self-Esteem distributions reveal Hispanics have the lowest self-esteem while whites and blacks have similar self-esteem distributions. White men and women scored highest on the AFQT test followed by their Hispanic and black counterparts. Comparing locus of control for men shows Hispanic men are the most external with white men being the most internal. Black women are the most external followed by Hispanic and white women.

2.6 Results

2.6.1 Ordinary Least Squares Regression Results

Tables 2.3 and 2.4 present ordinary least squares regressions for the pooled data and the time-averaged data. These results cover the ten survey years for 1991-2006, corresponding to the sample age beginning at 26-29 and ending at 41-45. The four specifications are estimated by gender (columns 1-4, men; columns 5-8, women). Columns 1 and 5 show the specification without cognitive or noncognitive measures; columns 2 and 6 add the cognitive measure; columns 3 and 7 add the noncognitive measure; and columns 4 and 8 add both cognitive and noncognitive measures. The first subtable uses the Rotter Scale while the second subtable uses the Rosenberg Scale.

In Table 2.3, first subtable (Rotter Scale) the black racial gap follows the same pattern discovered by Neal and Johnson (1996). Adding AFQT score dramatically reduces the magnitude of the negative coefficient on black men (column 2). For black women adding AFQT produces this effect too. The Hispanic coefficient on men falls and switches signs for women, qualitatively matching Neal and Johnson's results for Hispanics. Comparing the black and Hispanic coefficients for men in column 3 that includes the Rotter Score to column 1 shows very little change (about 1 percent reduction), so noncognitive skills cannot account for much of the wage gap. This agrees with Andrisani (1977) who could not find a large difference in the return to noncognitive skills between white and black men. The black coefficient for women after adding the Rotter Score (column 7) falls by about 2 percent, suggesting noncognitive skills account for a larger portion of the wage gap. After controlling for both sets of skills (columns 4 and 8), there still exists a return to cognitive skills for both men and women. For noncognitive skills men and women experience a significant return. The negative sign on the Rotter coefficient implies more external individuals receive lower wages. This return to internal individuals is consistent with Andrisani's (1977) analysis of earlier NLS data and Cebi's (2007) analysis of 2000 NLSY79 data.[22]

Noncognitive skills, as measured by the Rosenberg Self-Esteem Scale in the second subtable, slightly widen the black wage gap for men (column 3) (1 percent) but for women the gap widens more (column 7) (3 percent). When including self-esteem, the Hispanic gap for men falls by 2 percent, but there is no significant effect on the Hispanic gap for women. When including AFQT and self-esteem, men and women receive a positive, significant return to self-esteem. The return to women is higher. The positive coefficient means higher self-esteem translates to higher wages later in life. This positive relationship agrees with Goldsmith, Veum, and Darity (1997) who conduct their analysis with NLSY79 data from 1987. Like locus of control self-esteem seems to only significantly affect wages in a linear way.

Table 2.4 presents the same analysis using time-averaged data. The between estimator produces coefficient estimates that are qualitatively similar to the estimates in Table 2.3. Noncognitive skills measured by either the Rotter Scale or Rosenberg Scale change the black wage gap for men (column 3) by a small amount (1 percent reduction). Locus of control changes the wage gap for Hispanic men by a 1 percent reduction, but self-esteem reduces the gap by 2 percent. Noncognitive skills have differing effects for black women. Locus of control shrinks the gap by 2 percent, but self-esteem widens it by 2 percent. The Hispanic coefficient for women remains insignificant whether including

[22] Cebi (2007) defines the Rotter Scale, so a higher score implies a more internal individual. This analysis follows the NLSY documentation which defines the Rotter Scale in the opposite way, so the coefficients in this analysis will have the opposite sign as those reported in Cebi (2007).

locus of control or self-esteem. After controlling for AFQT and locus of control, men and women experience a significant return to locus of control. After controlling for AFQT and self-esteem, men and women experience a significant return to self-esteem. Women now face a lower return to self-esteem than men.

2.6.2 Quantile Regression Results

Tables 2.3 and 2.4 corroborate the main finding of Urzua (2008) that noncognitive skills can not account for much of the black-white wage gap for men measured at the mean of the wage distribution. Tables 2.5-2.8 extend the analysis presented in Tables 2.3 and 2.4 to examine effects at various quantiles of the wage distribution. Tables 2.5 and 2.6 show quantile regressions with locus of control and self-esteem, respectively, estimated on the pooled data. Figure 2.2 shows the change in the wage gap after adding noncognitive skills. It plots the difference in the black and Hispanic coefficients between Specification 1 and Specification 3. Cognitive skills are mainly responsible for the reduction of the wage gap in Table 2.5 for men and women. Locus of control does not greatly affect the magnitude of the wage gap for black and Hispanic men at any quantile (reduction of 1-2 percent); however, locus of control does affect the gap for black and Hispanic women. For black women locus of control accounts for a portion of the wage gap at quantiles 20-40 (about 2 percent reduction). This portion grows for quantiles 50-80 to about 5 percent at the 80th quantile and falls to about 3 percent at the 90th quantile. For Hispanic women locus of control accounts for a portion of the wage gap (about 2 percent reduction) in the upper quantiles (60 -90) too. Self-esteem (Table 2.6) has differing effects on the wage gap for black men across the wage distribution. The wage gap for black men falls by 1-2 percent when including self-esteem at some quantiles but mostly widens or does not change. For Hispanic men self-esteem closes the gap at most quantiles with the greatest reduction at the 10th and 70th quantiles (about 4 percent). The gap for black women grows by 4 percent at the 10th quantile, grows by 1-3 percent at quantiles 20-60, and grows by 4-5 percent at quantiles 70-90. Most of the estimates on the Hispanic coefficient for women are imprecisely estimated at quantiles at or below the median. For Hispanic women above the median the wage gap does not change or widens by 1-3 percent.

Tables 2.7 and 2.8 show quantile regressions estimated on the time-averaged data. Table 2.7 includes locus of control while Table 2.8 includes self-esteem. Still, cognitive skills are responsible for most of the reduction in the wage gap for men and women. Figure 2.3, like Figure 2.2, shows the change in the wage gap after adding noncognitive skills. At most quantiles locus of control influences the wage gap for black and Hispanic men usually by a reduction of 1-2 percent. For black women locus of control lowers the gap by 1-2 percent at lower quantiles (10-60) and by 3-6 percent at higher quantiles (70-90). Locus of control for Hispanic women has a negligible effect on the gap. Including self-esteem widens the wage gap at most quantiles for black men but lowers the gap at most quantiles for Hispanic men by 2-7 percent. Like black men black women experience a wider gap when including self-esteem at most quantiles with the largest differences at the upper quantiles (70-90) (3-10 percent). Self-esteem does not impact Hispanic women as most Hispanic coeffients are imprecisely estimated.

Figures 2.4 and 2.5 plot each quantile coefficient and its 95% confidence band for specifications 4 and 8 from Tables 2.5-2.8. These specifications control for both cognitive and noncognitive measures and give a sense of how the wage gaps and return to cognitive and noncognitive skills vary over the wage distribution. Figure 2.4 is based on pooled data, and Figure 2.5 is based on time-averaged data. Panels 1 (locus of control) and 2 (self-esteem) of Figure 2.4 show the specifications for locus of control by gender. These panels show differing effects on the black wage gap. The gap remains relatively flat and persists across the wage distribution for men and locus of control. Black women earn more than white women at the lowest quantile of the distribution (9.0 percent more) and continue to earn more until they earn less at the 90th quantile (3.6 percent less). Hispanic men earn less than their white counterparts at the 10th quantile (about 7 percent less) until the 60th quantile when they earn more for the remaining portions of the wage distribution.[23] Hispanic women, on the other hand, always earn more than white women across the wage distribution.[24] The profiles for AFQT and Rotter indicate a larger return to cognitive skills than noncognitive skills.[25] Using self-esteem as the noncognitive measure does not qualitatively change the trends. The wage gap persists for black men. Black women earn more than white women at the lower quantiles of the distribution and earn less at higher quantiles. Hispanic men face an upward sloping wage gap profile, and Hispanic women always earn more. Similarly, the return to cognitive skills exceeds the return to self-esteem across the entire distribution. These trends do not change with time-averaged data (Figure 2.5).[26]

2.6.3 Robustness

Tables 2.9-2.14 offer specifications that replace AFQT with components of AFQT and interact race with cognitive skills and noncognitive skills to allow for differing

[23] A joint hypothesis test on the equality of the Hispanic coefficient for men across quantiles is rejected at the 1 percent significance level.
[24] A joint hypothesis test on the equality of the Hispanic coefficient for women across quantiles is rejected at the 5 percent significance level.
[25] A joint hypothesis test on the equality of AFQT and Rotter coefficients for each gender across quantiles is rejected at the 1 percent significance level.
[26] A joint hypothesis test on the equality of Hispanic coefficients for men and women cannot be rejected at the 10 percent significance level.

returns by race. Tables 2.9-2.12 follow Table 2.3 in using pooled data but replace the composite AFQT score with each section that comprises it: word knowledge (Table 2.9), arithmetic reasoning (Table 2.10), paragraph comprehension (Table 2.11), and numerical operations (Table 2.12).[27] Overall, replacing the AFQT score with its components still dramatically reduces the magnitude of the negative coefficient on black men (column 2). The largest reduction occurs with the Word Knowledge score, and the smallest reduction occurs with the Numerical Operations score. None of these reductions are greater than the reduction when including the AFQT score. For black women using components of the AFQT produces large reductions too. The Hispanic coefficient on men falls and switches signs for women, qualitatively matching the results when using AFQT. After controlling for each section of the AFQT and locus of control, there still exists a return to locus of control for men and women. More external individuals earn less with the largest effect for Numerical Operations and Paragraph Comprehension. After controlling for each section of the AFQT and self-esteem, men and women with higher self-esteem receive higher wages later in life. The largest return occurs with Arithmetic Reasoning and Numerical Operations for men and women. The return to self-esteem for women still exceeds the return for men when using the components.

Tables 2.13 and 2.14 report specifications that interact race with AFQT and noncognitive skills to allow for varying returns by race. Like Table 2.3 both tables use pooled data. Table 2.13 shows the specification with locus of control, and Table 2.14 shows the specification with self-esteem. When controlling for both sets of skills, these specifications suggest no differential returns in men by race for locus of control or self-

[27] Each section score is normalized to have a mean of zero and standard deviation of one.

esteem (column 4). The specifications do suggest differential returns for women only in self-esteem. When controlling for AFQT and self-esteem, black and Hispanic women earn more than their white counterparts (column 8).

Tables 2.15, 2.16, and 2.17 report specifications that control for region of residence where individuals lived as a teen in 1979. Table 2.15 adds a region dummy variable for South to each specification and controls for the large black population residing in the South. Individuals who lived in the South as teens earn lower wages later in life, regardless of which noncognitive skill is examined. The black coefficient in column 1 falls by about 3 percent compared to the black coefficient in column 1 of Table 2.3 which does not include a region control. Comparing the AFQT and locus of control coefficients in Table 2.15 and Table 2.3 shows very little change. Though the magnitudes of the black and Hispanic coefficients are smaller than in Table 2.3, the qualitative results do not change after controlling for region. Tables 2.16 and 2.17 interact the South dummy variable with cognitive and noncognitive skills to allow for differing returns by the region where individuals lived as a teen. Both tables show a higher return to AFQT for women living in the South in 1979. Table 2.17 shows a nonlinear effect on self-esteem for women living in the South.

2.7 Conclusion and Future Work

This chapter investigates the role of noncognitive skills in explaining racial gaps in wages. Noncognitive skills are added to a parsimonious wage regression from Neal and Johnson (1996) to examine their effect on the wage gap. The analysis extends the wage gap and noncognitive skills literatures by studying the effect of noncognitive skills on wage gaps across the entire wage distribution. Using data from the National

Longitudinal Survey of Youth 1979 spanning 1991-2006 this chapter estimates wage regressions based on three estimators: a pooled estimator, a between estimator, and a quantile estimator. The wage regressions take advantage of the timing of when noncognitive skills and wages are measured. The wage regressions relate cognitive and noncognitive skills measured at the beginning of the NLSY before individuals enter the labor market or begin post secondary schooling to wages measured later in life. The various model specifications capture the separate and simultaneous effects of cognitive and noncognitive skills on the wage gap.

Model estimates based on the pooled and between estimators confirm the finding in the wage gap literature that cognitive skills consistently account for much of the male black-white wage gap measured at the mean of the wage distribution. These model estimates also confirm a finding in the noncognitive literature that noncognitive skills cannot account for the male black-white wage gap measured at the mean of the wage distribution. While the pooled and between estimators suggest significant returns exist to noncognitive skills even after controlling for cognitive skills, the rank ordering of these returns between genders differs by estimator. Quantile regressions of the specification controlling for cognitive and noncognitive skills have different implications on the black-white wage gap. The black-white male wage gap persists at all points of the wage distribution. The black-white female wage gap exists at the highest portion of the wage distribution. The Hispanic-white wage gap profiles also differ by gender. After controlling for cognitive and noncognitive skills, Hispanic men earn less than white men at lower quantiles but earn more at higher quantiles. After controlling for cognitive and noncognitive skills, Hispanic women earn more than white women at all quantiles. The return to cognitive skills is greater than the return to noncognitive skills at all quantiles after controlling for both sets of skills.

Noncognitive skills have generally been found to determine wage levels in the general population and across both genders (Heckman 2006). The noncognitive literature has emphasized the importance of connecting the development of these skills in early childhood to adult outcomes. In this context, the finding in this chapter that noncognitive skills cannot affect or close some racial wage gaps presents a puzzle to the noncognitive literature. On one hand, these skills are important for wage levels; on the other hand, they do not seem to be important for wage gaps. It is possible that the specific noncognitive skills examined do not dramatically differ by race as depicted by their density functions. It is also possible that other noncognitive skills that are more direct measures of work ethic and motivation may be better determinants of relative wages.

Future work relating noncognitive skills to the racial wage gap should address a few issues, most notably an adjustment for sample selection, alternative measures of noncognitive skills, and an investigation into using the AFQT as a proxy variable for premarket human capital. Chandra (2003) and Neal (2004) show ignoring labor force withdrawal biases estimates of the racial wage gap. Chandra (2003) and Neal (2004) implement variations of a matching estimator to impute wages for individuals with missing wage data. Similar procedures should be implemented on the wage regressions in this chapter to account for the influence of missing wages on the wage gap. Given the wide classification of a noncognitive skill, many alternative measures are available in the NLSY79. Weinberger (2008a and 2008b) and Rouse (2008) use sports and leadership participation as a measure of noncognitive skills which motivates an alternative measure as participation in extracurricular activities. In 1984 the NLSY79 asked questions about high school participation in sports, student government, student publications, performing arts, and clubs. Krueger and Schkade (2008) develop a measure of gregariousness based on time diary information to gauge how sociability impacts selection into jobs. In 1985 individuals in the NLSY79 self report the degree to which he or she is shy or outgoing as a measure of sociability. Carneiro, Heckman, and Masterov (2005) use the behavior problems index to measure noncognitive skills in children of the NLSY79 cohort. A series of questions in 1980 surveying school discipline problems related to suspension and expulsion could serve as a similar measure for the NLSY79 cohort. A set of questions about risk and impatience in the most recent survey, 2006, can be used to determine degree of risk aversion. Bollinger (2003) shows the measurement error associated with using the AFQT as a proxy variable for human capital accumulation may bias the racial coefficients in the specification. Future work should investigate the severity of this bias in this context where proxy variables for cognitive and noncognitive skills are used.

Table 2.1 Sample Summary Statistics

Variable	Mean	Std. Dev.	Min.	Max.
Entire Sample (n=25,085)				
Male	0.53	0.50	0	1
Female	0.47	0.50	0	1
Black	0.29	0.45	0	1
Hispanic	0.19	0.39	0	1
White	0.52	0.50	0	1
AFQT	0.01	0.82	-2.33	1.65
Rotter	0.18	1.00	-2.32	3.37
Rosenberg	0.18	0.94	-2.49	1.90
Age	33.67	4.72	26.00	43.00
Wage (2009 dollars)	16.72	10.20	1.00	74.75
Black Sample (n=7,311)				
Male	0.52	0.50	0	1
Female	0.48	0.50	0	1
AFQT	-0.45	0.68	-2.33	1.50
Rotter	0.29	1.06	-2.32	3.37
Rosenberg	-0.06	0.92	-2.49	1.90
Age	33.77	4.72	26.00	43.00
Wage (2009 dollars)	13.99	8.50	1.02	73.07
Hispanic Sample (n=4,822)				
Male	0.50	0.50	0	1
Female	0.50	0.50	0	1
AFQT	-0.21	0.74	-2.33	1.50
Rotter	0.23	0.99	-2.32	3.37
Rosenberg	-0.30	0.93	-2.49	1.90
Age	33.62	4.72	26.00	43.00
Wage (2009 dollars)	16.42	9.93	1.00	74.75
White Sample (n=12,952)				
Male	0.54	0.50	0	1
Female	0.46	0.50	0	1
AFQT	0.32	0.77	-2.33	1.65
Rotter	0.10	-0.95	2.32	3.37
Rosenberg	-0.20	-0.95	2.23	1.90
Age	33.63	4.72	26.00	43.00
Wage (2009 dollars)	18.37	10.82	1.00	73.53

Table 2.2 Sample Summary Statistics By Gender and Race

Men

Variable	Mean	Std. Dev.	Min.	Max.
Black (n=3,772)				
AFQT	-0.53	0.69	-2.33	1.40
Rotter	0.22	1.12	-2.32	3.37
Rosenberg	-0.09	0.94	-2.23	1.90
Age	33.61	4.69	26.00	42.00
Wage (2009 dollars)	14.83	9.13	1.02	73.07
Hispanic (n=2,424)				
AFQT	-0.27	0.78	-2.33	1.37
Rotter	0.25	1.01	-2.32	3.37
Rosenberg	-0.32	0.90	-2.49	1.90
Age	33.46	4.72	26.00	43.00
Wage (2009 dollars)	17.51	10.48	1.10	74.12
White (n=7,039)				
AFQT	0.26	0.81	-2.33	1.65
Rotter	0.09	0.95	-2.32	3.37
Rosenberg	-0.16	0.93	-2.23	1.90
Age	33.61	4.71	26.00	42.00
Wage (2009 dollars)	20.35	11.30	1.08	73.12

Women

Variable	Mean	Std. Dev.	Min.	Max.
Black (n=3,539)				
AFQT	-0.37	0.65	-2.33	1.50
Rotter	0.38	1.00	-2.32	3.37
Rosenberg	-0.04	0.89	-2.49	1.90
Age	33.93	4.73	26.00	43.00
Wage (2009 dollars)	13.10	7.66	1.08	67.60
Hispanic (n=2,398)				
AFQT	-0.14	0.68	-2.33	1.50
Rotter	0.21	0.97	-2.32	3.37
Rosenberg	-0.28	0.96	-2.23	1.90
Age	33.78	4.72	26.00	42.00
Wage (2009 dollars)	15.33	9.22	1.00	74.75
White (n=5,913)				
AFQT	0.39	0.71	-2.33	1.62
Rotter	0.12	0.96	-2.32	3.37
Rosenberg	-0.23	0.98	-2.23	1.90
Age	33.65	4.72	26.00	43.00
Wage (2009 dollars)	16.01	9.71	1.00	73.53

Table 2.3 Log Wage Regression Using Pooled Data with Locus of Control and Self-Esteem

	Men (1-4)				Women (5-8)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	-0.322***	-0.119***	-0.317***	-0.123***	-0.180***	0.0502*	-0.157***	0.0522**
	(0.024)	(0.025)	(0.024)	(0.025)	(0.027)	(0.027)	(0.026)	(0.026)
Hispanic	-0.155***	-0.0170	-0.148***	-0.0173	-0.0341	0.133***	-0.0250	0.131***
	(0.030)	(0.029)	(0.030)	(0.029)	(0.031)	(0.029)	(0.031)	(0.029)
Age	0.0305***	0.00289	0.0270**	0.00146	0.00244	-0.00493	-0.00533	-0.00876
	(0.011)	(0.011)	(0.011)	(0.011)	(0.013)	(0.011)	(0.013)	(0.011)
AFQT		0.259***		0.255***		0.294***		0.281***
		(0.014)		(0.014)		(0.016)		(0.016)
$AFQT^2$		0.0643***		0.0616***		0.0643***		0.0622***
		(0.014)		(0.014)		(0.015)		(0.015)
Rotter			-0.0522***	-0.0224**			-0.0812***	-0.0425***
			(0.011)	(0.010)			(0.014)	(0.012)
$Rotter^2$			0.00423	0.00737			0.00416	0.00355
			(0.0078)	(0.0069)			(0.0088)	(0.0079)
Observations	13163	13163	13137	13137	11819	11819	11765	11765
R^2	0.11	0.22	0.12	0.22	0.05	0.16	0.07	0.16

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	-0.322***	-0.119***	-0.331***	-0.139***	-0.180***	0.0502*	-0.207***	0.0160
	(0.024)	(0.025)	(0.024)	(0.025)	(0.027)	(0.027)	(0.026)	(0.027)
Hispanic	-0.155***	-0.0170	-0.139***	-0.0203	-0.0341	0.133***	-0.0281	0.123***
	(0.030)	(0.029)	(0.029)	(0.028)	(0.031)	(0.029)	(0.030)	(0.029)
Age	0.0305***	0.00289	0.0149	-0.00247	0.00244	-0.00493	-0.00517	-0.00841
	(0.011)	(0.011)	(0.011)	(0.011)	(0.013)	(0.011)	(0.012)	(0.011)
AFQT		0.259***		0.240***		0.294***		0.269***
		(0.014)		(0.015)		(0.016)		(0.017)
$AFQT^2$		0.0643***		0.0600***		0.0643***		0.0605***
		(0.014)		(0.013)		(0.015)		(0.014)
Rosenberg			0.111***	0.0527***			0.114***	0.0611***
			(0.011)	(0.011)			(0.012)	(0.012)
$Rosenberg^2$			-0.00928	-0.00516			-0.0177	-0.0117
			(0.010)	(0.0097)			(0.012)	(0.011)
Observations	13163	13163	13153	13153	11819	11819	11809	11809
R^2	0.11	0.22	0.14	0.23	0.05	0.16	0.08	0.17

Regressions include annual time dummy variables. Standard errors in parentheses correct for the longitudinal structure of the NLSY by accounting for repeated observations of individuals over time. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 2.4 Log Wage Regression Using Time-Averaged Data with Locus of Control and Self-Esteem

	Men (1-4)				Women (5-8)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	-0.367***	-0.151***	-0.361***	-0.157***	-0.163***	0.0624**	-0.139***	0.0659**
	(0.029)	(0.030)	(0.029)	(0.030)	(0.029)	(0.030)	(0.029)	(0.030)
Hispanic	-0.187***	-0.0333	-0.179***	-0.0343	-0.0424	0.120***	-0.0322	0.118***
	(0.034)	(0.033)	(0.034)	(0.033)	(0.035)	(0.034)	(0.034)	(0.034)
Age	0.0151*	0.00852	0.0148*	0.00849	-0.00773	-0.0154*	-0.0111	-0.0168*
	(0.0079)	(0.0074)	(0.0078)	(0.0074)	(0.0091)	(0.0086)	(0.0091)	(0.0086)
AFQT		0.279***		0.270***		0.287***		0.274***
		(0.016)		(0.016)		(0.020)		(0.020)
AFQT^2		0.0643***		0.0604***		0.0574***		0.0548***
		(0.015)		(0.015)		(0.018)		(0.018)
Rotter			-0.0662***	-0.0353***			-0.0777***	-0.0406***
			(0.013)	(0.012)			(0.015)	(0.014)
Rotter^2			0.00230	0.00510			-0.00333	-0.00505
			(0.010)	(0.0097)			(0.0100)	(0.0092)
Observations	1675	1675	1671	1671	1588	1588	1580	1580
R^2	0.09	0.23	0.11	0.23	0.02	0.15	0.04	0.16

	Men (1-4)				Women (5-8)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	-0.367***	-0.151***	-0.378***	-0.182***	-0.163***	0.0624**	-0.185***	0.0343
	(0.029)	(0.030)	(0.028)	(0.031)	(0.029)	(0.030)	(0.029)	(0.031)
Hispanic	-0.187***	-0.0333	-0.163***	-0.0374	-0.0424	0.120***	-0.0346	0.111***
	(0.034)	(0.033)	(0.033)	(0.033)	(0.035)	(0.034)	(0.034)	(0.034)
Age	0.0151*	0.00852	0.0120	0.00757	-0.00773	-0.0154*	-0.0109	-0.0165*
	(0.0079)	(0.0074)	(0.0077)	(0.0074)	(0.0091)	(0.0086)	(0.0090)	(0.0086)
AFQT		0.279***		0.248***		0.287***		0.266***
		(0.016)		(0.017)		(0.020)		(0.020)
AFQT^2		0.0643***		0.0566***		0.0574***		0.0538***
		(0.015)		(0.015)		(0.018)		(0.018)
Rosenberg			0.138***	0.0767***			0.105***	0.0514***
			(0.013)	(0.013)			(0.014)	(0.014)
Rosenberg^2			-0.0114	-0.00655			-0.0158	-0.0112
			(0.012)	(0.011)			(0.014)	(0.013)
Observations	1675	1675	1674	1674	1588	1588	1587	1587
R^2	0.09	0.23	0.15	0.25	0.02	0.15	0.06	0.16

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 2.5 Quantile Wage Regression Using Pooled Data with Locus of Control

Quantile Wage Regression With Rotter Locus of Control, 1991-2006 (10th Quantile)

	Men (1-4)				Women (5-8)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	-0.238***	-0.116***	-0.246***	-0.121***	-0.0572***	0.0882***	-0.0619***	0.0904***
	(0.0159)	(0.0149)	(0.0139)	(0.0189)	(0.0163)	(0.0197)	(0.0150)	(0.0181)
Hispanic	-0.156***	-0.0661***	-0.164***	-0.0697***	-0.0179	0.0935***	-0.0356*	0.0902***
	(0.0221)	(0.0180)	(0.0186)	(0.0195)	(0.0196)	(0.0201)	(0.0184)	(0.0224)
Age	0.0157*	-0.00477	0.0134*	-0.00651	-0.00153	-0.00826	-0.00644	-0.00809
	(0.00898)	(0.00796)	(0.00737)	(0.00884)	(0.00636)	(0.0103)	(0.00775)	(0.00992)
AFQT		0.187***		0.186***		0.211***		0.206***
		(0.0118)		(0.0133)		(0.0146)		(0.0145)
$AFQT^2$		0.0232**		0.0224*		0.0449***		0.0478***
		(0.0111)		(0.0120)		(0.0130)		(0.0119)
Rotter			-0.0317***	-0.00511			-0.0445***	-0.0245***
			(0.00681)	(0.00818)			(0.00594)	(0.00916)
$Rotter^2$			0.000910	0.00484			0.00664	0.00865
			(0.00454)	(0.00445)			(0.00424)	(0.00591)
Observations	13163	13163	13137	13137	11819	11819	11765	11765

Regressions include annual time dummy variables. Standard errors in parentheses are based on 100 bootstrap replications. *** p<0.01, ** p<0.05, * p<0.1

Quantile Wage Regression With Rotter Locus of Control, 1991-2006 (20th Quantile)

	Men (1-4)				Women (5-8)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	-0.322***	-0.144***	-0.316***	-0.149***	-0.138***	0.0622***	-0.124***	0.0665***
	(0.0144)	(0.0130)	(0.0154)	(0.0148)	(0.0130)	(0.0160)	(0.0121)	(0.0141)
Hispanic	-0.176***	-0.0793***	-0.177***	-0.0822***	-0.0242	0.110***	-0.0334*	0.105***
	(0.0193)	(0.0202)	(0.0176)	(0.0205)	(0.0198)	(0.0152)	(0.0180)	(0.0181)
Age	0.0249***	-0.000352	0.0273***	-0.000181	0.00374	-0.0119*	0.00246	-0.0122*
	(0.00760)	(0.00683)	(0.00718)	(0.00673)	(0.00695)	(0.00687)	(0.00761)	(0.00646)
AFQT		0.249***		0.246***		0.257***		0.253***
		(0.00891)		(0.0102)		(0.0101)		(0.0108)
$AFQT^2$		0.0505***		0.0500***		0.0680***		0.0709***
		(0.00942)		(0.00943)		(0.00798)		(0.00729)
Rotter			-0.0280***	-0.0156**			-0.0618***	-0.0350***
			(0.00769)	(0.00687)			(0.00835)	(0.00796)
$Rotter^2$			-0.000995	0.0123***			0.0111*	0.0108**
			(0.00443)	(0.00394)			(0.00566)	(0.00421)
Observations	13163	13163	13137	13137	11819	11819	11765	11765

Regressions include annual time dummy variables. Standard errors in parentheses are based on 100 bootstrap replications. *** p<0.01, ** p<0.05, * p<0.1

Table 2.5 (continued)

Quantile Wage Regression With Rotter Locus of Control, 1991-2006 (30th Quantile)

	Men (1-4)				Women (5-8)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	-0.371***	-0.157***	-0.375***	-0.163***	-0.167***	0.0705***	-0.148***	0.0736***
	(0.0122)	(0.0151)	(0.0147)	(0.0136)	(0.0129)	(0.0135)	(0.0131)	(0.0138)
Hispanic	-0.205***	-0.0551***	-0.209***	-0.0614***	-0.0100	0.148***	-0.00749	0.149***
	(0.0164)	(0.0174)	(0.0184)	(0.0172)	(0.0192)	(0.0142)	(0.0270)	(0.0161)
Age	0.0263***	0.00379	0.0226***	0.00275	-0.00444	-0.0164***	-0.0121*	-0.0162**
	(0.00651)	(0.00596)	(0.00672)	(0.00588)	(0.00711)	(0.00558)	(0.00707)	(0.00654)
AFQT		0.272***		0.269***		0.301***		0.289***
		(0.00876)		(0.00808)		(0.00901)		(0.00972)
$AFQT^2$		0.0536***		0.0529***		0.0770***		0.0752***
		(0.00835)		(0.00741)		(0.00610)		(0.00844)
Rotter			-0.0419***	-0.0222***			-0.0719***	-0.0341***
			(0.00649)	(0.00543)			(0.00828)	(0.00752)
$Rotter^2$			0.00843*	0.00986***			0.0113*	0.00303
			(0.00470)	(0.00350)			(0.00627)	(0.00422)
Observations	13163	13163	13137	13137	11819	11819	11765	11765

Regressions include annual time dummy variables. Standard errors in parentheses are based on 100 bootstrap replications. *** p<0.01, ** p<0.05, * p<0.1

Quantile Wage Regression With Rotter Locus of Control, 1991-2006 (40th Quantile)

	Men (1-4)				Women (5-8)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	-0.374***	-0.137***	-0.369***	-0.144***	-0.170***	0.0611***	-0.150***	0.0623***
	(0.0127)	(0.0126)	(0.0124)	(0.0149)	(0.0136)	(0.0135)	(0.0148)	(0.0172)
Hispanic	-0.170***	-0.0238	-0.163***	-0.0182	0.00447	0.144***	0.0105	0.143***
	(0.0158)	(0.0152)	(0.0185)	(0.0122)	(0.0165)	(0.0154)	(0.0176)	(0.0160)
Age	0.0345***	0.00700	0.0316***	0.00692	-0.00447	-0.00777	-0.00883	-0.0131*
	(0.00688)	(0.00631)	(0.00635)	(0.00516)	(0.00693)	(0.00646)	(0.00703)	(0.00715)
AFQT		0.283***		0.278***		0.319***		0.306***
		(0.00716)		(0.00682)		(0.00864)		(0.00878)
$AFQT^2$		0.0553***		0.0493***		0.0771***		0.0747***
		(0.00613)		(0.00701)		(0.00660)		(0.00631)
Rotter			-0.0499***	-0.0226***			-0.0833***	-0.0365***
			(0.00634)	(0.00590)			(0.00905)	(0.00819)
$Rotter^2$			0.0123***	0.0106**			0.0104*	0.00325
			(0.00428)	(0.00465)			(0.00584)	(0.00511)
Observations	13163	13163	13137	13137	11819	11819	11765	11765

Regressions include annual time dummy variables. Standard errors in parentheses are based on 100 bootstrap replications. *** p<0.01, ** p<0.05, * p<0.1

Table 2.5 (continued)

Quantile Wage Regression With Rotter Locus of Control, 1991-2006 (50th Quantile)

	Men (1-4)				Women (5-8)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	-0.362***	-0.132***	-0.356***	-0.132***	-0.194***	0.0482***	-0.168***	0.0551***
	(0.0114)	(0.0132)	(0.0150)	(0.0138)	(0.0151)	(0.0143)	(0.0151)	(0.0136)
Hispanic	-0.172***	-0.00920	-0.167***	-0.00751	-0.0192	0.135***	-0.0115	0.134***
	(0.0149)	(0.0146)	(0.0244)	(0.0149)	(0.0165)	(0.0153)	(0.0180)	(0.0148)
Age	0.0440***	0.0146***	0.0382***	0.0119*	-0	0.00126	-0.0152**	-0.00316
	(0.00580)	(0.00548)	(0.0119)	(0.00632)	(0.00712)	(0.00587)	(0.00773)	(0.00557)
AFQT		0.283***		0.282***		0.331***		0.317***
		(0.00678)		(0.00724)		(0.00748)		(0.00864)
$AFQT^2$		0.0556***		0.0504***		0.0754***		0.0711***
		(0.00583)		(0.00691)		(0.00688)		(0.00723)
Rotter			-0.0516***	-0.0247***			-0.0998***	-0.0453***
			(0.00677)	(0.00494)			(0.00764)	(0.00703)
$Rotter^2$			0.00825	0.0106***			0.0149***	0.00454
			(0.0100)	(0.00362)			(0.00469)	(0.00455)
Observations	13163	13163	13137	13137	11819	11819	11765	11765

Regressions include annual time dummy variables. Standard errors in parentheses are based on 100 bootstrap replications. *** p<0.01, ** p<0.05, * p<0.1

Quantile Wage Regression With Rotter Locus of Control, 1991-2006 (60th Quantile)

	Men (1-4)				Women (5-8)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	-0.360***	-0.114***	-0.348***	-0.118***	-0.241***	0.0536***	-0.202***	0.0622***
	(0.0151)	(0.0136)	(0.0196)	(0.0134)	(0.0174)	(0.0127)	(0.0167)	(0.0198)
Hispanic	-0.166***	-0.000908	-0.159***	0.00463	-0.0583***	0.130***	-0.0347**	0.124***
	(0.0133)	(0.0147)	(0.0286)	(0.0143)	(0.0191)	(0.0173)	(0.0155)	(0.0157)
Age	0.0410***	0.0152**	0.0392**	0.0113**	0.00122	-0.000428	-0.00458	0.000442
	(0.00518)	(0.00597)	(0.0190)	(0.00525)	(0.00825)	(0.00646)	(0.00772)	(0.00628)
AFQT		0.287***		0.286***		0.341***		0.328***
		(0.00706)		(0.00744)		(0.00777)		(0.0155)
$AFQT^2$		0.0628***		0.0581***		0.0713***		0.0687***
		(0.00688)		(0.00743)		(0.00696)		(0.0157)
Rotter			-0.0580***	-0.0264***			-0.112***	-0.0575***
			(0.00660)	(0.00536)			(0.00771)	(0.0123)
$Rotter^2$			0.00814	0.0117***			0.0160***	0.0113*
			(0.00839)	(0.00373)			(0.00436)	(0.00669)
Observations	13163	13163	13137	13137	11819	11819	11765	11765

Regressions include annual time dummy variables. Standard errors in parentheses are based on 100 bootstrap replications. *** p<0.01, ** p<0.05, * p<0.1

Table 2.5 (continued)

Quantile Wage Regression With Rotter Locus of Control, 1991-2006 (70th Quantile)

	Men (1-4)				Women (5-8)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	-0.354***	-0.0972***	-0.333***	-0.0972***	-0.264***	0.0379**	-0.218***	0.0471***
	(0.0153)	(0.0128)	(0.0153)	(0.0142)	(0.0148)	(0.0174)	(0.0148)	(0.0169)
Hispanic	-0.173***	0.0189	-0.158***	0.0197	-0.0808***	0.136***	-0.0667***	0.138***
	(0.0163)	(0.0145)	(0.0161)	(0.0157)	(0.0174)	(0.0186)	(0.0167)	(0.0203)
Age	0.0426***	0.0114**	0.0369***	0.0104	0.0157**	0.0106	0.00911	-0.00112
	(0.00674)	(0.00570)	(0.00658)	(0.00676)	(0.00725)	(0.00763)	(0.00725)	(0.00686)
AFQT		0.296***		0.291***		0.335***		0.323***
		(0.00699)		(0.00743)		(0.00905)		(0.00820)
$AFQT^2$		0.0697***		0.0676***		0.0617***		0.0624***
		(0.00755)		(0.00661)		(0.00740)		(0.00771)
Rotter			-0.0656***	-0.0243***			-0.0956***	-0.0694***
			(0.00673)	(0.00555)			(0.00830)	(0.00677)
$Rotter^2$			0.00352	0.00790*			0.00404	0.0112***
			(0.00387)	(0.00415)			(0.00549)	(0.00430)
Observations	13163	13163	13137	13137	11819	11819	11765	11765

Regressions include annual time dummy variables. Standard errors in parentheses are based on 100 bootstrap replications. *** p<0.01, ** p<0.05, * p<0.1

Quantile Wage Regression With Rotter Locus of Control, 1991-2006 (80th Quantile)

	Men (1-4)				Women (5-8)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	-0.336***	-0.0899***	-0.328***	-0.0921***	-0.267***	0.0132	-0.216***	0.0231
	(0.0146)	(0.0147)	(0.0147)	(0.0142)	(0.0222)	(0.0194)	(0.0182)	(0.0181)
Hispanic	-0.163***	0.0230	-0.146***	0.0278	-0.0905***	0.129***	-0.0631***	0.125***
	(0.0208)	(0.0170)	(0.0214)	(0.0171)	(0.0192)	(0.0177)	(0.0225)	(0.0199)
Age	0.0404***	0.00294	0.0329***	0.00540	0.0136	0.00981	0.00356	0.00599
	(0.00655)	(0.00643)	(0.00755)	(0.00759)	(0.00905)	(0.00763)	(0.00733)	(0.00755)
AFQT		0.287***		0.280***		0.332***		0.316***
		(0.00866)		(0.00785)		(0.0114)		(0.0103)
$AFQT^2$		0.0829***		0.0788***		0.0569***		0.0583***
		(0.00900)		(0.00908)		(0.00903)		(0.00806)
Rotter			-0.0706***	-0.0256***			-0.0927***	-0.0592***
			(0.00785)	(0.00532)			(0.00683)	(0.00631)
$Rotter^2$			0.00189	0.00605**			-0.00203	0.00153
			(0.00516)	(0.00244)			(0.00461)	(0.00402)
Observations	13163	13163	13137	13137	11819	11819	11765	11765

Regressions include annual time dummy variables. Standard errors in parentheses are based on 100 bootstrap replications. *** p<0.01, ** p<0.05, * p<0.1

Table 2.5 (continued)

Quantile Wage Regression With Rotter Locus of Control, 1991-2006 (90th Quantile)

	Men (1-4)				Women (5-8)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	-0.300***	-0.0946***	-0.291***	-0.101***	-0.237***	-0.0324*	-0.211***	-0.0364*
	(0.0191)	(0.0240)	(0.0202)	(0.0199)	(0.0168)	(0.0182)	(0.0168)	(0.0200)
Hispanic	-0.0947***	0.0353	-0.0834***	0.0404*	-0.0616***	0.0956***	-0.0383	0.0944***
	(0.0216)	(0.0254)	(0.0197)	(0.0212)	(0.0232)	(0.0189)	(0.0274)	(0.0232)
Age	0.0408***	-0.00684	0.0320***	-0.000755	0.0210**	0.00946	0.0106	0.000189
	(0.0101)	(0.00777)	(0.00808)	(0.00925)	(0.00825)	(0.00920)	(0.0117)	(0.0108)
AFQT		0.243***		0.233***		0.300***		0.272***
		(0.0114)		(0.00935)		(0.0132)		(0.0169)
$AFQT^2$		0.0932***		0.0841***		0.0748***		0.0696***
		(0.0112)		(0.00999)		(0.0125)		(0.0141)
Rotter			-0.0787***	-0.0415***			-0.0821***	-0.0392***
			(0.00841)	(0.0101)			(0.00959)	(0.00780)
$Rotter^2$			0.00501	0.00271			-0.0105**	-0.00908*
			(0.00555)	(0.00621)			(0.00525)	(0.00478)
Observations	13163	13163	13137	13137	11819	11819	11765	11765

Regressions include annual time dummy variables. Standard errors in parentheses are based on 100 bootstrap replications. *** p<0.01, ** p<0.05, * p<0.1

Table 2.6 Quantile Wage Regression Using Pooled Data with Self-Esteem

Quantile Wage Regression With Rosenberg Self-Esteem, 1991-2006 (10th Quantile)

	Men (1-4)				Women (5-8)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	-0.238***	-0.116***	-0.258***	-0.137***	-0.0572***	0.0882***	-0.0963***	0.0648***
	(0.0166)	(0.0167)	(0.0172)	(0.0170)	(0.0160)	(0.0180)	(0.0134)	(0.0198)
Hispanic	-0.156***	-0.0661***	-0.118***	-0.0772***	-0.0179	0.0935***	-0.00648	0.0958***
	(0.0204)	(0.0180)	(0.0225)	(0.0155)	(0.0194)	(0.0216)	(0.0154)	(0.0224)
Age	0.0157*	-0.00477	0.00432	-0.00890	-0.00153	-0.00826	-0.00415	-0.00694
	(0.00868)	(0.00851)	(0.00860)	(0.00800)	(0.00747)	(0.00880)	(0.00749)	(0.00824)
AFQT		0.187***		0.177***		0.211***		0.196***
		(0.0124)		(0.0121)		(0.0137)		(0.0139)
$AFQT^2$		0.0232**		0.0228**		0.0449***		0.0431***
		(0.0106)		(0.0105)		(0.0122)		(0.0124)
Rosenberg			0.0782***	0.0297***			0.0638***	0.0400***
			(0.00781)	(0.00879)			(0.00766)	(0.00701)
$Rosenberg^2$			-0.00584	-0.00350			-0.0266***	-0.0172**
			(0.00846)	(0.00764)			(0.00800)	(0.00837)
Observations	13163	13163	13153	13153	11819	11819	11809	11809

Regressions include annual time dummy variables. Standard errors in parentheses are based on 100 bootstrap replications. *** p<0.01, ** p<0.05, * p<0.1

Quantile Wage Regression With Rosenberg Self-Esteem, 1991-2006 (20th Quantile)

	Men (1-4)				Women (5-8)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	-0.322***	-0.144***	-0.326***	-0.161***	-0.138***	0.0622***	-0.146***	0.0377***
	(0.0115)	(0.0140)	(0.0142)	(0.0149)	(0.00987)	(0.0164)	(0.0119)	(0.0142)
Hispanic	-0.176***	-0.0793***	-0.168***	-0.0854***	-0.0242	0.110***	0.00370	0.104***
	(0.0191)	(0.0186)	(0.0177)	(0.0206)	(0.0193)	(0.0146)	(0.0166)	(0.0175)
Age	0.0249***	-0.000352	0.0216***	-0.00151	0.00374	-0.0119*	-0.00447	-0.0143**
	(0.00720)	(0.00769)	(0.00770)	(0.00793)	(0.00713)	(0.00640)	(0.00683)	(0.00660)
AFQT		0.249***		0.230***		0.257***		0.239***
		(0.00982)		(0.0110)		(0.0111)		(0.0101)
$AFQT^2$		0.0505***		0.0453***		0.0680***		0.0682***
		(0.00894)		(0.00914)		(0.00938)		(0.00739)
Rosenberg			0.0886***	0.0393***			0.0866***	0.0491***
			(0.00748)	(0.00714)			(0.00704)	(0.00715)
$Rosenberg^2$			0.00307	0.00285			-0.0209***	-0.0149**
			(0.00704)	(0.00645)			(0.00650)	(0.00670)
Observations	13163	13163	13153	13153	11819	11819	11809	11809

Regressions include annual time dummy variables. Standard errors in parentheses are based on 100 bootstrap replications. *** p<0.01, ** p<0.05, * p<0.1

Table 2.6 (continued)

Quantile Wage Regression With Rosenberg Self-Esteem, 1991-2006 (30th Quantile)

	Men (1-4)				Women (5-8)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	-0.371***	-0.157***	-0.374***	-0.173***	-0.167***	0.0705***	-0.174***	0.0450***
	(0.0129)	(0.0147)	(0.0123)	(0.0158)	(0.0122)	(0.0118)	(0.0128)	(0.0152)
Hispanic	-0.205***	-0.0551***	-0.174***	-0.0572***	-0.0100	0.148***	0.000640	0.146***
	(0.0169)	(0.0184)	(0.0155)	(0.0204)	(0.0184)	(0.0137)	(0.0162)	(0.0159)
Age	0.0263***	0.00379	0.0184***	0.00128	-0.00444	-0.0164***	-0.00805	-0.0154**
	(0.00711)	(0.00655)	(0.00507)	(0.00607)	(0.00779)	(0.00634)	(0.00529)	(0.00629)
AFQT		0.272***		0.256***		0.301***		0.279***
		(0.00843)		(0.00994)		(0.00790)		(0.0102)
$AFQT^2$		0.0536***		0.0523***		0.0770***		0.0744***
		(0.00789)		(0.00710)		(0.00641)		(0.00703)
Rosenberg			0.104***	0.0410***			0.110***	0.0578***
			(0.00611)	(0.00634)			(0.00672)	(0.00669)
$Rosenberg^2$			0.00615	0.00241			-0.0259***	-0.0138**
			(0.00681)	(0.00546)			(0.00558)	(0.00542)
Observations	13163	13163	13153	13153	11819	11819	11809	11809

Regressions include annual time dummy variables. Standard errors in parentheses are based on 100 bootstrap replications. *** p<0.01, ** p<0.05, * p<0.1

Quantile Wage Regression With Rosenberg Self-Esteem, 1991-2006 (40th Quantile)

	Men (1-4)				Women (5-8)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	-0.374***	-0.137***	-0.386***	-0.150***	-0.170***	0.0611***	-0.195***	0.0371***
	(0.0111)	(0.0157)	(0.0139)	(0.0166)	(0.0153)	(0.0128)	(0.0148)	(0.0142)
Hispanic	-0.170***	-0.0238	-0.159***	-0.0181	0.00447	0.144***	0.0157	0.144***
	(0.0166)	(0.0153)	(0.0167)	(0.0147)	(0.0152)	(0.0160)	(0.0160)	(0.0146)
Age	0.0345***	0.00700	0.0238***	0.00399	-0.00447	-0.00777	-0.00746	-0.00929
	(0.00573)	(0.00530)	(0.00602)	(0.00584)	(0.00637)	(0.00538)	(0.00683)	(0.00574)
AFQT		0.283***		0.268***		0.319***		0.294***
		(0.00622)		(0.00882)		(0.00803)		(0.00845)
$AFQT^2$		0.0553***		0.0529***		0.0771***		0.0784***
		(0.00635)		(0.00718)		(0.00666)		(0.00596)
Rosenberg			0.105***	0.0416***			0.123***	0.0675***
			(0.00556)	(0.00583)			(0.00650)	(0.00540)
$Rosenberg^2$			-0.000434	0.00136			-0.0207***	-0.0159***
			(0.00558)	(0.00513)			(0.00706)	(0.00591)
Observations	13163	13163	13153	13153	11819	11819	11809	11809

Regressions include annual time dummy variables. Standard errors in parentheses are based on 100 bootstrap replications. *** p<0.01, ** p<0.05, * p<0.1

Table 2.6 (continued)

Quantile Wage Regression With Rosenberg Self-Esteem, 1991-2006 (50th Quantile)

	Men (1-4)				Women (5-8)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	-0.362***	-0.132***	-0.378***	-0.142***	-0.194***	0.0482***	-0.227***	0.0154
	(0.0155)	(0.0137)	(0.0171)	(0.0131)	(0.0157)	(0.0134)	(0.0140)	(0.0150)
Hispanic	-0.172***	-0.00920	-0.163***	-0.0122	-0.0192	0.135***	-0.0232	0.127***
	(0.0133)	(0.0148)	(0.0297)	(0.0127)	(0.0163)	(0.0183)	(0.0151)	(0.0142)
Age	0.0440***	0.0146***	0.0242*	0.00688	-0	0.00126	-0.00315	-0.00404
	(0.00657)	(0.00540)	(0.0136)	(0.00556)	(0.00757)	(0.00636)	(0.00697)	(0.00568)
AFQT		0.283***		0.266***		0.331***		0.307***
		(0.00679)		(0.00744)		(0.00798)		(0.00800)
$AFQT^2$		0.0556***		0.0518***		0.0754***		0.0723***
		(0.00553)		(0.00559)		(0.00734)		(0.00572)
Rosenberg			0.117***	0.0473***			0.140***	0.0693***
			(0.00599)	(0.00540)			(0.00712)	(0.00643)
$Rosenberg^2$			-0.00121	-0.00595			-0.0214***	-0.0103
			(0.00919)	(0.00526)			(0.00600)	(0.00627)
Observations	13163	13163	13153	13153	11819	11819	11809	11809

Regressions include annual time dummy variables. Standard errors in parentheses are based on 100 bootstrap replications. *** p<0.01, ** p<0.05, * p<0.1

Quantile Wage Regression With Rosenberg Self-Esteem, 1991-2006 (60th Quantile)

	Men (1-4)				Women (5-8)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	-0.360***	-0.114***	-0.354***	-0.134***	-0.241***	0.0536***	-0.265***	-0.00900
	(0.0140)	(0.0134)	(0.0376)	(0.0132)	(0.0173)	(0.0143)	(0.0155)	(0.0151)
Hispanic	-0.166***	-0.000908	-0.145***	-0.00382	-0.0583***	0.130***	-0.0549***	0.101***
	(0.0126)	(0.0142)	(0.0329)	(0.0131)	(0.0212)	(0.0171)	(0.0180)	(0.0169)
Age	0.0410***	0.0152**	0.0274	0.00758	0.00122	-0.000428	0.00107	0.00253
	(0.00532)	(0.00665)	(0.225)	(0.00567)	(0.00772)	(0.00618)	(0.00711)	(0.00653)
AFQT		0.287***		0.267***		0.341***		0.302***
		(0.00708)		(0.00825)		(0.00756)		(0.00911)
$AFQT^2$		0.0628***		0.0603***		0.0713***		0.0635***
		(0.00620)		(0.00640)		(0.00613)		(0.00666)
Rosenberg			0.123	0.0534***			0.151***	0.0816***
			(0.652)	(0.00720)			(0.00714)	(0.00706)
$Rosenberg^2$			-0.00115	-0.00166			-0.0166**	-0.0107*
			(0.806)	(0.00582)			(0.00739)	(0.00638)
Observations	13163	13163	13153	13153	11819	11819	11809	11809

Regressions include annual time dummy variables. Standard errors in parentheses are based on 100 bootstrap replications. *** p<0.01, ** p<0.05, * p<0.1

Table 2.6 (continued)

Quantile Wage Regression With Rosenberg Self-Esteem, 1991-2006 (70th Quantile)

	Men (1-4)				Women (5-8)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	-0.354***	-0.0972***	-0.348***	-0.125***	-0.264***	0.0379**	-0.303***	-0.0317*
	(0.0172)	(0.0124)	(0.0160)	(0.0155)	(0.0160)	(0.0176)	(0.0176)	(0.0168)
Hispanic	-0.173***	0.0189	-0.138***	0.00617	-0.0808***	0.136***	-0.0795***	0.102***
	(0.0164)	(0.0148)	(0.0203)	(0.0148)	(0.0167)	(0.0174)	(0.0182)	(0.0181)
Age	0.0426***	0.0114*	0.0216***	0.00529	0.0157**	0.0106	0.00606	0.00516
	(0.00632)	(0.00597)	(0.00737)	(0.00718)	(0.00665)	(0.00753)	(0.00809)	(0.00661)
AFQT		0.296***		0.269***		0.335***		0.292***
		(0.00725)		(0.00704)		(0.00771)		(0.0101)
AFQT^2		0.0697***		0.0640***		0.0617***		0.0554***
		(0.00761)		(0.00860)		(0.00860)		(0.00759)
Rosenberg			0.128***	0.0605***			0.152***	0.0934***
			(0.00624)	(0.00666)			(0.00710)	(0.00843)
Rosenberg^2			-0.0127**	-0.000641			-0.0221***	-0.00645
			(0.00643)	(0.00556)			(0.00771)	(0.00635)
Observations	13163	13163	13153	13153	11819	11819	11809	11809

Regressions include annual time dummy variables. Standard errors in parentheses are based on 100 bootstrap replications. *** p<0.01, ** p<0.05, * p<0.1

Quantile Wage Regression With Rosenberg Self-Esteem, 1991-2006 (80th Quantile)

	Men (1-4)				Women (5-8)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	-0.336***	-0.0899***	-0.344***	-0.115***	-0.267***	0.0132	-0.320***	-0.0547***
	(0.0146)	(0.0143)	(0.0158)	(0.0125)	(0.0225)	(0.0179)	(0.0147)	(0.0176)
Hispanic	-0.163***	0.0230	-0.150***	0.0182	-0.0905***	0.129***	-0.0970***	0.0917***
	(0.0211)	(0.0180)	(0.0166)	(0.0161)	(0.0186)	(0.0165)	(0.0193)	(0.0175)
Age	0.0404***	0.00294	0.0181***	-0.00551	0.0136	0.00981	0.00692	-0.00210
	(0.00623)	(0.00736)	(0.00691)	(0.00746)	(0.00997)	(0.00833)	(0.00768)	(0.00720)
AFQT		0.287***		0.259***		0.332***		0.290***
		(0.00791)		(0.00786)		(0.0104)		(0.0119)
AFQT^2		0.0829***		0.0749***		0.0569***		0.0486***
		(0.00933)		(0.00872)		(0.00819)		(0.00804)
Rosenberg			0.131***	0.0657***			0.138***	0.0879***
			(0.00729)	(0.00653)			(0.00838)	(0.00799)
Rosenberg^2			-0.0191***	-0.00979			-0.00868	-0.00434
			(0.00606)	(0.00599)			(0.00732)	(0.00606)
Observations	13163	13163	13153	13153	11819	11819	11809	11809

Regressions include annual time dummy variables. Standard errors in parentheses are based on 100 bootstrap replications. *** p<0.01, ** p<0.05, * p<0.1

Table 2.6 (continued)

Quantile Wage Regression With Rosenberg Self-Esteem, 1991-2006 (90th Quantile)

	Men (1-4)				Women (5-8)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	-0.300***	-0.0946***	-0.290***	-0.121***	-0.237***	-0.0324*	-0.290***	-0.0696***
	(0.0193)	(0.0183)	(0.0200)	(0.0204)	(0.0165)	(0.0181)	(0.0212)	(0.0210)
Hispanic	-0.0947***	0.0353	-0.106***	0.0226	-0.0616***	0.0956***	-0.0937***	0.0911***
	(0.0216)	(0.0229)	(0.0229)	(0.0242)	(0.0226)	(0.0191)	(0.0172)	(0.0228)
Age	0.0408***	-0.00684	0.0176**	-0.0186*	0.0210**	0.00946	-1.59e-05	0.00213
	(0.00947)	(0.00827)	(0.00883)	(0.0102)	(0.00972)	(0.00882)	(0.0105)	(0.00953)
AFQT		0.243***		0.220***		0.300***		0.278***
		(0.0104)		(0.0117)		(0.0153)		(0.0130)
$AFQT^2$		0.0932***		0.0868***		0.0748***		0.0736***
		(0.0114)		(0.00943)		(0.0149)		(0.0132)
Rosenberg			0.124***	0.0796***			0.123***	0.0564***
			(0.00667)	(0.00898)			(0.00873)	(0.00851)
$Rosenberg^2$			-0.0366***	-0.0196***			0.00314	-0.00411
			(0.00598)	(0.00679)			(0.00923)	(0.00907)
Observations	13163	13163	13153	13153	11819	11819	11809	11809

Regressions include annual time dummy variables. Standard errors in parentheses are based on 100 bootstrap replications. *** p<0.01, ** p<0.05, * p<0.1

Table 2.7 Quantile Wage Regression Using Time-Averaged Data with Locus of Control

Quantile Wage Regression With Rotter Locus of Control, Between Estimator 1991-2006 (10th Quantile)

	Men (1-4)				Women (5-8)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	-0.282***	-0.162***	-0.273***	-0.199***	-0.0635	0.0893	-0.0842	0.0675
	(0.0365)	(0.0418)	(0.0351)	(0.0445)	(0.0414)	(0.0583)	(0.0553)	(0.0632)
Hispanic	-0.189***	-0.117**	-0.192***	-0.117**	-0.0742	-0.0184	-0.0685	-0.0194
	(0.0568)	(0.0491)	(0.0488)	(0.0518)	(0.0787)	(0.0829)	(0.0656)	(0.0817)
Age	0.0125	0.0120	0.0129	0.0164	0.00426	0.00873	0.00187	0.00495
	(0.0156)	(0.0133)	(0.0152)	(0.0132)	(0.0179)	(0.0139)	(0.0200)	(0.0169)
AFQT		0.209***		0.207***		0.177***		0.171***
		(0.0319)		(0.0358)		(0.0482)		(0.0518)
AFQT^2		0.0109		0.0271		-0.0129		-0.00576
		(0.0316)		(0.0303)		(0.0631)		(0.0721)
Rotter			-0.0597***	-0.0402**			-0.0406	-0.0281
			(0.0219)	(0.0174)			(0.0291)	(0.0341)
Rotter^2			0.00425	0.00923			4.20e-05	0.00306
			(0.0191)	(0.0136)			(0.0225)	(0.0172)
Observations	1675	1675	1671	1671	1588	1588	1580	1580

Standard errors in parentheses are based on 100 bootstrap replications. *** p<0.01, ** p<0.05, * p<0.1

Quantile Wage Regression With Rotter Locus of Control, Between Estimator 1991-2006 (20th Quantile)

	Men				Women			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	-0.319***	-0.149***	-0.321***	-0.160***	-0.0748**	0.0947*	-0.0585	0.0715
	(0.0461)	(0.0324)	(0.0382)	(0.0388)	(0.0379)	(0.0539)	(0.0397)	(0.0465)
Hispanic	-0.209***	-0.0678	-0.203***	-0.0858	-0.0173	0.109*	-0.0103	0.128*
	(0.0520)	(0.0539)	(0.0417)	(0.0615)	(0.0475)	(0.0631)	(0.0495)	(0.0652)
Age	0.0164	0.0135*	0.0162	0.0124	-0.00267	-0.0143	-0.00403	-0.0110
	(0.0107)	(0.00815)	(0.0104)	(0.0107)	(0.0146)	(0.0131)	(0.0152)	(0.0135)
AFQT		0.260***		0.252***		0.251***		0.230***
		(0.0195)		(0.0186)		(0.0340)		(0.0366)
AFQT^2		0.0444**		0.0498***		0.0596**		0.0579**
		(0.0174)		(0.0184)		(0.0283)		(0.0230)
Rotter			-0.0561***	-0.0250			-0.0528**	-0.0326**
			(0.0156)	(0.0156)			(0.0206)	(0.0166)
Rotter^2			0.00249	0.00848			-0.00566	0.00102
			(0.0110)	(0.0116)			(0.0136)	(0.0103)
Observations	1675	1675	1671	1671	1588	1588	1580	1580

Standard errors in parentheses are based on 100 bootstrap replications. *** p<0.01, ** p<0.05, * p<0.1

Table 2.7 (continued)

Quantile Wage Regression With Rotter Locus of Control, Between Estimator 1991-2006 (30th Quantile)

	Men (1-4)				Women (5-8)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	-0.382***	-0.142***	-0.373***	-0.158***	-0.117***	0.0790**	-0.110***	0.0639**
	(0.0326)	(0.0354)	(0.0337)	(0.0364)	(0.0348)	(0.0343)	(0.0282)	(0.0309)
Hispanic	-0.220***	-0.0680	-0.215***	-0.0737*	0.0107	0.144***	-0.00146	0.136***
	(0.0500)	(0.0464)	(0.0517)	(0.0397)	(0.0486)	(0.0404)	(0.0372)	(0.0341)
Age	0.0143	0.0218**	0.0111	0.0174**	-0.0134	-0.0175	-0.0115	-0.0219*
	(0.00970)	(0.00899)	(0.0103)	(0.00853)	(0.0130)	(0.0115)	(0.0114)	(0.0116)
AFQT		0.271***		0.260***		0.263***		0.245***
		(0.0243)		(0.0252)		(0.0215)		(0.0218)
$AFQT^2$		0.0482***		0.0418**		0.0651***		0.0601***
		(0.0186)		(0.0171)		(0.0209)		(0.0192)
Rotter			-0.0531***	-0.0295**			-0.0607***	-0.0356**
			(0.0154)	(0.0141)			(0.0184)	(0.0156)
$Rotter^2$			0.00353	0.00180			-0.00409	-0.00559
			(0.0132)	(0.0113)			(0.0156)	(0.0106)
Observations	1675	1675	1671	1671	1588	1588	1580	1580

Standard errors in parentheses are based on 100 bootstrap replications. *** p<0.01, ** p<0.05, * p<0.1

Quantile Wage Regression With Rotter Locus of Control, Between Estimator 1991-2006 (40th Quantile)

	Men (1-4)				Women (5-8)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	-0.412***	-0.144***	-0.420***	-0.144***	-0.142***	0.0646**	-0.120***	0.0717**
	(0.0302)	(0.0395)	(0.0264)	(0.0424)	(0.0312)	(0.0317)	(0.0291)	(0.0286)
Hispanic	-0.214***	-0.0515	-0.206***	-0.0547	0.000327	0.127***	0.00292	0.130***
	(0.0392)	(0.0409)	(0.0402)	(0.0396)	(0.0386)	(0.0284)	(0.0449)	(0.0378)
Age	0.00963	0.0132	0.00931	0.0126	-0.0145	-0.0181**	-0.0134	-0.0201**
	(0.00854)	(0.00984)	(0.00790)	(0.00839)	(0.00909)	(0.00862)	(0.0106)	(0.00992)
AFQT		0.289***		0.285***		0.277***		0.271***
		(0.0239)		(0.0224)		(0.0211)		(0.0196)
$AFQT^2$		0.0479**		0.0474**		0.0727***		0.0683***
		(0.0210)		(0.0185)		(0.0147)		(0.0133)
Rotter			-0.0516***	-0.0258			-0.0545***	-0.0325*
			(0.0134)	(0.0164)			(0.0176)	(0.0170)
$Rotter^2$			0.00482	0.0117			-0.00177	-0.00447
			(0.0119)	(0.0132)			(0.0153)	(0.0118)
Observations	1675	1675	1671	1671	1588	1588	1580	1580

Standard errors in parentheses are based on 100 bootstrap replications. *** p<0.01, ** p<0.05, * p<0.1

Table 2.7 (continued)

Quantile Wage Regression With Rotter Locus of Control, Between Estimator 1991-2006 (50th Quantile)

	Men (1-4)				Women (5-8)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	-0.415***	-0.131***	-0.404***	-0.128***	-0.159***	0.0469	-0.137***	0.0745**
	(0.0300)	(0.0385)	(0.0340)	(0.0414)	(0.0336)	(0.0356)	(0.0330)	(0.0376)
Hispanic	-0.188***	-0.0311	-0.180***	-0.0313	0.0283	0.121***	0.0204	0.118***
	(0.0376)	(0.0383)	(0.0319)	(0.0392)	(0.0435)	(0.0356)	(0.0441)	(0.0348)
Age	0.0134*	0.0119	0.0113	0.0104	-0.0171	-0.0199*	-0.0203*	-0.0167
	(0.00701)	(0.00826)	(0.00834)	(0.00805)	(0.0108)	(0.0107)	(0.0116)	(0.0124)
AFQT		0.296***		0.304***		0.309***		0.299***
		(0.0204)		(0.0202)		(0.0249)		(0.0223)
$AFQT^2$		0.0579***		0.0545***		0.0743***		0.0712***
		(0.0173)		(0.0153)		(0.0154)		(0.0166)
Rotter			-0.0491***	-0.0267*			-0.0836***	-0.0398**
			(0.0132)	(0.0161)			(0.0174)	(0.0172)
$Rotter^2$			0.0121	0.0169			0.00460	-0.0113
			(0.0112)	(0.0120)			(0.0118)	(0.0112)
Observations	1675	1675	1671	1671	1588	1588	1580	1580

Standard errors in parentheses are based on 100 bootstrap replications. *** p<0.01, ** p<0.05, * p<0.1

Quantile Wage Regression With Rotter Locus of Control, Between Estimator 1991-2006 (60th Quantile)

	Men (1-4)				Women (5-8)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	-0.415***	-0.127***	-0.384***	-0.147***	-0.196***	0.0586*	-0.183***	0.0805***
	(0.0375)	(0.0395)	(0.0452)	(0.0394)	(0.0354)	(0.0352)	(0.0467)	(0.0306)
Hispanic	-0.190***	0.0124	-0.178***	-0.00840	-0.00351	0.176***	-0.00537	0.165***
	(0.0369)	(0.0497)	(0.0377)	(0.0487)	(0.0453)	(0.0576)	(0.0532)	(0.0474)
Age	0.0186**	0.0128*	0.0167*	0.0118	-0.0173*	-0.0161	-0.0218*	-0.0198
	(0.00903)	(0.00770)	(0.00903)	(0.00808)	(0.0105)	(0.0126)	(0.0113)	(0.0131)
AFQT		0.306***		0.295***		0.336***		0.331***
		(0.0173)		(0.0216)		(0.0180)		(0.0271)
$AFQT^2$		0.0559***		0.0552***		0.0883***		0.0824***
		(0.0162)		(0.0177)		(0.0175)		(0.0230)
Rotter			-0.0605***	-0.0181			-0.0976***	-0.0400
			(0.0169)	(0.0150)			(0.0297)	(0.0324)
$Rotter^2$			0.0149	0.0102			0.0124	-0.00764
			(0.0133)	(0.0108)			(0.0164)	(0.0135)
Observations	1675	1675	1671	1671	1588	1588	1580	1580

Standard errors in parentheses are based on 100 bootstrap replications. *** p<0.01, ** p<0.05, * p<0.1

Table 2.7 (continued)

Quantile Wage Regression With Rotter Locus of Control, Between Estimator 1991-2006 (70th Quantile)

	Men (1-4)				Women (5-8)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	-0.376***	-0.116***	-0.374***	-0.125***	-0.275***	0.0564	-0.211***	0.0580
	(0.0396)	(0.0357)	(0.0428)	(0.0371)	(0.0501)	(0.0343)	(0.0488)	(0.0377)
Hispanic	-0.205***	0.0173	-0.168***	0.00968	-0.0649	0.182***	-0.0465	0.157***
	(0.0393)	(0.0340)	(0.0462)	(0.0384)	(0.0541)	(0.0419)	(0.0429)	(0.0449)
Age	0.0124	0.0126	0.00777	0.0153*	-0.0118	-0.0270**	-0.0184*	-0.0288***
	(0.00979)	(0.00929)	(0.00908)	(0.00842)	(0.0119)	(0.0107)	(0.0108)	(0.0101)
AFQT		0.306***		0.299***		0.354***		0.347***
		(0.0220)		(0.0195)		(0.0202)		(0.0232)
$AFQT^2$		0.0630***		0.0582***		0.0916***		0.0855***
		(0.0151)		(0.0193)		(0.0188)		(0.0190)
Rotter			-0.0723***	-0.0263**			-0.108***	-0.0486**
			(0.0151)	(0.0127)			(0.0181)	(0.0221)
$Rotter^2$			0.0121	0.00885			0.0145	-0.000528
			(0.00936)	(0.00885)			(0.0114)	(0.0109)
Observations	1675	1675	1671	1671	1588	1588	1580	1580

Standard errors in parentheses are based on 100 bootstrap replications. *** p<0.01, ** p<0.05, * p<0.1

Quantile Wage Regression With Rotter Locus of Control, Between Estimator 1991-2006 (80th Quantile)

	Men (1-4)				Women (5-8)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	-0.330***	-0.122***	-0.366***	-0.129***	-0.272***	0.0523	-0.242***	0.0748*
	(0.0437)	(0.0306)	(0.0394)	(0.0370)	(0.0422)	(0.0436)	(0.0546)	(0.0383)
Hispanic	-0.170***	0.0244	-0.172***	0.00331	-0.107**	0.148***	-0.113**	0.174***
	(0.0572)	(0.0476)	(0.0503)	(0.0465)	(0.0447)	(0.0470)	(0.0459)	(0.0480)
Age	0.0108	0.00326	0.0143	0.00467	-0.0112	-0.0251**	-0.0130	-0.0238**
	(0.0120)	(0.00874)	(0.0101)	(0.0102)	(0.0114)	(0.0108)	(0.0135)	(0.0121)
AFQT		0.302***		0.296***		0.376***		0.345***
		(0.0211)		(0.0210)		(0.0246)		(0.0225)
$AFQT^2$		0.0792***		0.0733***		0.0766***		0.0768***
		(0.0218)		(0.0200)		(0.0221)		(0.0242)
Rotter			-0.0746***	-0.0361**			-0.0911***	-0.0615***
			(0.0191)	(0.0176)			(0.0192)	(0.0171)
$Rotter^2$			0.00342	0.0119			-0.00487	-0.00669
			(0.0119)	(0.00799)			(0.00986)	(0.0105)
Observations	1675	1675	1671	1671	1588	1588	1580	1580

Standard errors in parentheses are based on 100 bootstrap replications. *** p<0.01, ** p<0.05, * p<0.1

Table 2.7 (continued)

Quantile Wage Regression With Rotter Locus of Control, Between Estimator 1991-2006 (90th Quantile)

	Men (1-4)				Women (5-8)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	-0.352***	-0.116**	-0.348***	-0.115**	-0.210***	0.0282	-0.179***	0.0366
	(0.0490)	(0.0503)	(0.0583)	(0.0501)	(0.0554)	(0.0509)	(0.0555)	(0.0643)
Hispanic	-0.140**	0.0267	-0.122	0.0344	-0.132**	0.146***	-0.0875	0.106
	(0.0668)	(0.0578)	(0.0873)	(0.0490)	(0.0560)	(0.0474)	(0.0738)	(0.0657)
Age	0.0209*	-0.00372	0.0134	0.000833	-0.000870	-0.0143	-0.0152	-0.0160
	(0.0127)	(0.0136)	(0.0119)	(0.0128)	(0.0138)	(0.00966)	(0.0151)	(0.0119)
AFQT		0.277***		0.274***		0.362***		0.338***
		(0.0278)		(0.0258)		(0.0343)		(0.0349)
$AFQT^2$		0.108***		0.0946***		0.0754**		0.0588*
		(0.0290)		(0.0259)		(0.0362)		(0.0344)
Rotter			-0.0634***	-0.0481**			-0.0705**	-0.0507**
			(0.0243)	(0.0214)			(0.0289)	(0.0243)
$Rotter^2$			0.00508	0.00244			-0.0207	-0.0160
			(0.0176)	(0.0117)			(0.0139)	(0.0122)
Observations	1675	1675	1671	1671	1588	1588	1580	1580

Standard errors in parentheses are based on 100 bootstrap replications. *** p<0.01, ** p<0.05, * p<0.1

Table 2.8 Quantile Wage Regression Using Time-Averaged Data with Self-Esteem

Quantile Wage Regression With Rosenberg Self-Esteem, Between Estimator 1991-2006 (10th Quantile)

	Men (1-4)				Women (5-8)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	-0.282***	-0.162***	-0.291***	-0.199***	-0.0635	0.0893	-0.0588	0.0616
	(0.0386)	(0.0401)	(0.0489)	(0.0499)	(0.0444)	(0.0647)	(0.0506)	(0.0659)
Hispanic	-0.189***	-0.117**	-0.116**	-0.109*	-0.0742	-0.0184	-0.0367	0.0119
	(0.0567)	(0.0516)	(0.0520)	(0.0578)	(0.0662)	(0.0952)	(0.0866)	(0.0887)
Age	0.0125	0.0120	0.00675	0.0127	0.00426	0.00873	0.000963	0.00590
	(0.0159)	(0.0117)	(0.0137)	(0.0150)	(0.0154)	(0.0165)	(0.0186)	(0.0160)
AFQT		0.209***		0.185***		0.177***		0.167***
		(0.0307)		(0.0455)		(0.0576)		(0.0528)
$AFQT^2$		0.0109		0.0164		-0.0129		0.00135
		(0.0288)		(0.0326)		(0.0613)		(0.0573)
Rosenberg			0.120***	0.0510**			0.0415	0.0200
			(0.0209)	(0.0247)			(0.0282)	(0.0231)
$Rosenberg^2$			-0.00584	-0.0131			-0.0365	-0.0300
			(0.0185)	(0.0206)			(0.0240)	(0.0234)
Observations	1675	1675	1674	1674	1588	1588	1587	1587

Standard errors in parentheses are based on 100 bootstrap replications. *** p<0.01, ** p<0.05, * p<0.1

Quantile Wage Regression With Rosenberg Self-Esteem, Between Estimator 1991-2006 (20th Quantile)

	Men (1-4)				Women (5-8)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	-0.319***	-0.149***	-0.336***	-0.168***	-0.0748*	0.0947**	-0.0917***	0.0329
	(0.0439)	(0.0350)	(0.0379)	(0.0380)	(0.0390)	(0.0476)	(0.0319)	(0.0449)
Hispanic	-0.209***	-0.0678	-0.171***	-0.0693	-0.0173	0.109*	0.0400	0.0856
	(0.0569)	(0.0475)	(0.0477)	(0.0492)	(0.0452)	(0.0627)	(0.0502)	(0.0562)
Age	0.0164	0.0135	0.0182*	0.0144	-0.00267	-0.0143	-0.00282	-0.00502
	(0.0111)	(0.00838)	(0.00949)	(0.00979)	(0.0130)	(0.0143)	(0.0131)	(0.0108)
AFQT		0.260***		0.249***		0.251***		0.210***
		(0.0204)		(0.0253)		(0.0332)		(0.0284)
$AFQT^2$		0.0444***		0.0554***		0.0596**		0.0547***
		(0.0171)		(0.0171)		(0.0234)		(0.0200)
Rosenberg			0.113***	0.0496**			0.0703***	0.0484***
			(0.0207)	(0.0198)			(0.0175)	(0.0149)
$Rosenberg^2$			0.000855	-0.00124			-0.0297**	-0.0308**
			(0.0196)	(0.0175)			(0.0143)	(0.0141)
Observations	1675	1675	1674	1674	1588	1588	1587	1587

Standard errors in parentheses are based on 100 bootstrap replications. *** p<0.01, ** p<0.05, * p<0.1

Table 2.8 (continued)

Quantile Wage Regression With Rosenberg Self-Esteem, Between Estimator 1991-2006 (30th Quantile)

	Men (1-4)				Women (5-8)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	-0.382***	-0.142***	-0.396***	-0.178***	-0.117***	0.0790**	-0.156***	0.0450
	(0.0356)	(0.0390)	(0.0305)	(0.0378)	(0.0319)	(0.0330)	(0.0268)	(0.0339)
Hispanic	-0.220***	-0.0680	-0.180***	-0.0497	0.0107	0.144***	0.00478	0.137***
	(0.0496)	(0.0440)	(0.0452)	(0.0425)	(0.0450)	(0.0356)	(0.0494)	(0.0427)
Age	0.0143	0.0218**	0.0179	0.0164	-0.0134	-0.0175	-0.0156	-0.0134
	(0.0104)	(0.00918)	(0.0112)	(0.0106)	(0.0123)	(0.0122)	(0.0112)	(0.0109)
AFQT		0.271***		0.257***		0.263***		0.247***
		(0.0236)		(0.0245)		(0.0228)		(0.0236)
$AFQT^2$		0.0482**		0.0522**		0.0651***		0.0612***
		(0.0214)		(0.0230)		(0.0176)		(0.0217)
Rosenberg			0.116***	0.0534***			0.0849***	0.0429***
			(0.0154)	(0.0121)			(0.0168)	(0.0138)
$Rosenberg^2$			-0.00970	-0.0112			-0.0352**	-0.0326**
			(0.0171)	(0.0127)			(0.0143)	(0.0143)
Observations	1675	1675	1674	1674	1588	1588	1587	1587

Standard errors in parentheses are based on 100 bootstrap replications. *** p<0.01, ** p<0.05, * p<0.1

Quantile Wage Regression With Rosenberg Self-Esteem, Between Estimator 1991-2006 (40th Quantile)

	Men (1-4)				Women (5-8)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	-0.412***	-0.144***	-0.395***	-0.160***	-0.142***	0.0646**	-0.154***	0.0364
	(0.0301)	(0.0417)	(0.0284)	(0.0371)	(0.0312)	(0.0322)	(0.0339)	(0.0327)
Hispanic	-0.214***	-0.0515	-0.185***	-0.0249	0.000327	0.127***	0.0249	0.132***
	(0.0491)	(0.0388)	(0.0413)	(0.0435)	(0.0403)	(0.0317)	(0.0382)	(0.0329)
Age	0.00963	0.0132	0.00986	0.0123	-0.0145	-0.0181**	-0.0161	-0.0186**
	(0.00981)	(0.00922)	(0.00798)	(0.00848)	(0.00954)	(0.00800)	(0.0119)	(0.00896)
AFQT		0.289***		0.269***		0.277***		0.261***
		(0.0257)		(0.0208)		(0.0218)		(0.0208)
$AFQT^2$		0.0479**		0.0458***		0.0727***		0.0700***
		(0.0196)		(0.0156)		(0.0165)		(0.0163)
Rosenberg			0.122***	0.0650***			0.0983***	0.0445**
			(0.0166)	(0.0147)			(0.0185)	(0.0190)
$Rosenberg^2$			-0.00468	-0.00323			-0.0270	-0.0228
			(0.0148)	(0.0145)			(0.0177)	(0.0168)
Observations	1675	1675	1674	1674	1588	1588	1587	1587

Standard errors in parentheses are based on 100 bootstrap replications. *** p<0.01, ** p<0.05, * p<0.1

Table 2.8 (continued)

Quantile Wage Regression With Rosenberg Self-Esteem, Between Estimator 1991-2006 (50th Quantile)

	Men (1-4)				Women (5-8)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	-0.415***	-0.131***	-0.440***	-0.156***	-0.159***	0.0469	-0.155***	0.0706*
	(0.0277)	(0.0369)	(0.0298)	(0.0359)	(0.0367)	(0.0403)	(0.0321)	(0.0383)
Hispanic	-0.188***	-0.0311	-0.184***	-0.0258	0.0283	0.121***	0.0399	0.134***
	(0.0372)	(0.0416)	(0.0383)	(0.0367)	(0.0484)	(0.0392)	(0.0352)	(0.0359)
Age	0.0134	0.0119	0.0124	0.00878	-0.0171*	-0.0199**	-0.0214*	-0.0170
	(0.00841)	(0.00872)	(0.0109)	(0.00757)	(0.0101)	(0.00976)	(0.0115)	(0.0108)
AFQT		0.296***		0.275***		0.309***		0.301***
		(0.0183)		(0.0219)		(0.0246)		(0.0287)
$AFQT^2$		0.0579***		0.0564***		0.0743***		0.0817***
		(0.0152)		(0.0156)		(0.0188)		(0.0206)
Rosenberg			0.119***	0.0634***			0.121***	0.0488***
			(0.0152)	(0.0152)			(0.0173)	(0.0178)
$Rosenberg^2$			0.00285	0.00815			-0.0294**	-0.0151
			(0.0146)	(0.0160)			(0.0150)	(0.0160)
Observations	1675	1675	1674	1674	1588	1588	1587	1587

Standard errors in parentheses are based on 100 bootstrap replications. *** p<0.01, ** p<0.05, * p<0.1

Quantile Wage Regression With Rosenberg Self-Esteem, Between Estimator 1991-2006 (60th Quantile)

	Men (1-4)				Women (5-8)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	-0.415***	-0.127***	-0.403***	-0.199***	-0.196***	0.0586**	-0.213***	0.00926
	(0.0364)	(0.0353)	(0.0313)	(0.0416)	(0.0314)	(0.0283)	(0.0391)	(0.0287)
Hispanic	-0.190***	0.0124	-0.171***	-0.0284	-0.00351	0.176***	0.00480	0.142***
	(0.0411)	(0.0493)	(0.0460)	(0.0440)	(0.0414)	(0.0499)	(0.0417)	(0.0381)
Age	0.0186*	0.0128	0.0148	0.00792	-0.0173	-0.0161	-0.0185*	-0.0159
	(0.00981)	(0.00807)	(0.00925)	(0.00789)	(0.0119)	(0.0112)	(0.0102)	(0.0109)
AFQT		0.306***		0.266***		0.336***		0.302***
		(0.0185)		(0.0211)		(0.0194)		(0.0229)
$AFQT^2$		0.0559***		0.0518***		0.0883***		0.0857***
		(0.0168)		(0.0169)		(0.0175)		(0.0162)
Rosenberg			0.136***	0.0728***			0.143***	0.0732***
			(0.0174)	(0.0205)			(0.0172)	(0.0168)
$Rosenberg^2$			0.0119	0.00200			-0.0279*	-0.0196
			(0.0179)	(0.0154)			(0.0151)	(0.0158)
Observations	1675	1675	1674	1674	1588	1588	1587	1587

Standard errors in parentheses are based on 100 bootstrap replications. *** p<0.01, ** p<0.05, * p<0.1

Table 2.8 (continued)

Quantile Wage Regression With Rosenberg Self-Esteem, Between Estimator 1991-2006 (70th Quantile)

	Men (1-4)				Women (5-8)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	-0.376***	-0.116***	-0.387***	-0.143***	-0.275***	0.0564	-0.302***	0.00467
	(0.0426)	(0.0364)	(0.0353)	(0.0335)	(0.0486)	(0.0370)	(0.0413)	(0.0425)
Hispanic	-0.205***	0.0173	-0.160***	-0.0120	-0.0649	0.182***	-0.0689	0.140***
	(0.0504)	(0.0339)	(0.0464)	(0.0369)	(0.0490)	(0.0436)	(0.0441)	(0.0386)
Age	0.0124	0.0126	0.0143	0.0183**	-0.0118	-0.0270**	-0.00893	-0.0241*
	(0.0114)	(0.00803)	(0.00874)	(0.00723)	(0.0126)	(0.0110)	(0.00952)	(0.0128)
AFQT		0.306***		0.281***		0.354***		0.324***
		(0.0197)		(0.0213)		(0.0190)		(0.0260)
$AFQT^2$		0.0630***		0.0533***		0.0916***		0.0813***
		(0.0188)		(0.0153)		(0.0192)		(0.0161)
Rosenberg			0.143***	0.0606***			0.159***	0.0660***
			(0.0163)	(0.0141)			(0.0194)	(0.0178)
$Rosenberg^2$			0.00609	-0.00212			-0.00481	-0.00583
			(0.0154)	(0.0119)			(0.0218)	(0.0186)
Observations	1675	1675	1674	1674	1588	1588	1587	1587

Standard errors in parentheses are based on 100 bootstrap replications. *** p<0.01, ** p<0.05, * p<0.1

Quantile Wage Regression With Rosenberg Self-Esteem, Between Estimator 1991-2006 (80th Quantile)

	Men (1-4)				Women (5-8)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	-0.330***	-0.122***	-0.346***	-0.147***	-0.272***	0.0523	-0.342***	-0.0205
	(0.0417)	(0.0326)	(0.0396)	(0.0314)	(0.0468)	(0.0398)	(0.0534)	(0.0482)
Hispanic	-0.170***	0.0244	-0.150***	0.0116	-0.107**	0.148***	-0.131***	0.123**
	(0.0449)	(0.0495)	(0.0475)	(0.0403)	(0.0440)	(0.0404)	(0.0447)	(0.0496)
Age	0.0108	0.00326	0.0110	0.00531	-0.0112	-0.0251**	-0.0179	-0.0212**
	(0.0126)	(0.00952)	(0.00938)	(0.0107)	(0.0104)	(0.0114)	(0.0116)	(0.0105)
AFQT		0.302***		0.274***		0.376***		0.338***
		(0.0189)		(0.0198)		(0.0233)		(0.0309)
$AFQT^2$		0.0792***		0.0655***		0.0766***		0.0757***
		(0.0222)		(0.0230)		(0.0214)		(0.0232)
Rosenberg			0.161***	0.0762***			0.156***	0.0810***
			(0.0167)	(0.0198)			(0.0201)	(0.0211)
$Rosenberg^2$			-0.0105	-0.00697			-0.0201	0.00570
			(0.0160)	(0.0137)			(0.0210)	(0.0204)
Observations	1675	1675	1674	1674	1588	1588	1587	1587

Standard errors in parentheses are based on 100 bootstrap replications. *** p<0.01, ** p<0.05, * p<0.1

Table 2.8 (continued)

Quantile Wage Regression With Rosenberg Self-Esteem, Between Estimator 1991-2006 (90th Quantile)

	Men (1-4)				Women (5-8)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	-0.352***	-0.116**	-0.359***	-0.161***	-0.210***	0.0282	-0.311***	-0.0306
	(0.0528)	(0.0489)	(0.0384)	(0.0511)	(0.0533)	(0.0552)	(0.0551)	(0.0698)
Hispanic	-0.140**	0.0267	-0.169***	-0.00113	-0.132***	0.146***	-0.136***	0.0923
	(0.0689)	(0.0544)	(0.0564)	(0.0468)	(0.0410)	(0.0531)	(0.0492)	(0.0600)
Age	0.0209**	-0.00372	0.0258**	-0.00156	-0.000870	-0.0143	-0.0102	-0.0173
	(0.0103)	(0.0114)	(0.0108)	(0.0129)	(0.0155)	(0.0112)	(0.0136)	(0.0137)
AFQT		0.277***		0.231***		0.362***		0.320***
		(0.0286)		(0.0299)		(0.0359)		(0.0379)
$AFQT^2$		0.108***		0.0827***		0.0754**		0.0688*
		(0.0265)		(0.0215)		(0.0311)		(0.0361)
Rosenberg			0.153***	0.0980***			0.135***	0.0695***
			(0.0156)	(0.0166)			(0.0249)	(0.0236)
$Rosenberg^2$			-0.0327*	-0.0207			0.00976	0.00500
			(0.0168)	(0.0193)			(0.0240)	(0.0225)
Observations	1675	1675	1674	1674	1588	1588	1587	1587

Standard errors in parentheses are based on 100 bootstrap replications. *** p<0.01, ** p<0.05, * p<0.1

Table 2.9 Log Wage Regression Using Pooled Data with Word Knowledge, Locus of Control, and Self-Esteem

	Men (1-4)				Women (5-8)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	-0.322***	-0.130***	-0.317***	-0.135***	-0.180***	0.0243	-0.157***	0.0261
	(0.0241)	(0.0257)	(0.0243)	(0.0258)	(0.0266)	(0.0273)	(0.0265)	(0.0271)
Hispanic	-0.155***	-0.0376	-0.148***	-0.0379	-0.0341	0.111***	-0.0250	0.109***
	(0.0298)	(0.0290)	(0.0299)	(0.0290)	(0.0310)	(0.0303)	(0.0307)	(0.0302)
Age	0.0305***	0.0135	0.0270**	0.0116	0.00244	-0.00840	-0.00533	-0.0124
	(0.0115)	(0.0107)	(0.0115)	(0.0108)	(0.0127)	(0.0116)	(0.0126)	(0.0116)
Word Know		0.203***		0.198***		0.222***		0.211***
		(0.0148)		(0.0149)		(0.0164)		(0.0164)
Word Know^2		0.0120		0.00971		-0.0122		-0.0144
		(0.0104)		(0.0104)		(0.0122)		(0.0125)
Rotter			-0.0522***	-0.0280***			-0.0812***	-0.0472***
			(0.0113)	(0.0107)			(0.0136)	(0.0129)
Rotter^2			0.00423	0.00979			0.00416	0.00709
			(0.00777)	(0.00731)			(0.00880)	(0.00815)
Observations	13163	12981	13137	12955	11819	11640	11765	11586
R^2	0.111	0.200	0.119	0.203	0.048	0.142	0.065	0.146

	Men (1-4)				Women (5-8)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	-0.322***	-0.130***	-0.331***	-0.154***	-0.180***	0.0243	-0.207***	-0.0144
	(0.0241)	(0.0257)	(0.0235)	(0.0260)	(0.0266)	(0.0273)	(0.0261)	(0.0279)
Hispanic	-0.155***	-0.0376	-0.139***	-0.0413	-0.0341	0.111***	-0.0281	0.0985***
	(0.0298)	(0.0290)	(0.0287)	(0.0287)	(0.0310)	(0.0303)	(0.0295)	(0.0297)
Age	0.0305***	0.0135	0.0149	0.00718	0.00244	-0.00840	-0.00517	-0.0116
	(0.0115)	(0.0107)	(0.0111)	(0.0106)	(0.0127)	(0.0116)	(0.0121)	(0.0113)
Word Know		0.203***		0.184***		0.222***		0.198***
		(0.0148)		(0.0153)		(0.0164)		(0.0166)
Word Know^2		0.0120		0.0119		-0.0122		-0.0113
		(0.0104)		(0.0105)		(0.0122)		(0.0120)
Rosenberg			0.111***	0.0567***			0.114***	0.0663***
			(0.0111)	(0.0111)			(0.0125)	(0.0125)
Rosenberg^2			-0.00928	-0.00535			-0.0177	-0.0136
			(0.0104)	(0.00991)			(0.0121)	(0.0117)
Observations	13163	12981	13153	12971	11819	11640	11809	11630
R^2	0.111	0.200	0.144	0.209	0.048	0.142	0.084	0.153

Regressions include annual time dummy variables. Standard errors in parentheses correct for the longitudinal structure of the NLSY by accounting for repeated observations of individuals over time. *** p<0.01, ** p<0.05, * p<0.1

Table 2.10 Log Wage Regression Using Pooled Data with Arithmetic Reasoning, Locus of Control, and Self-Esteem

	Men (1-4)				Women (5-8)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	-0.322***	-0.143***	-0.317***	-0.147***	-0.180***	0.0105	-0.157***	0.0141
	(0.0241)	(0.0252)	(0.0243)	(0.0254)	(0.0266)	(0.0285)	(0.0265)	(0.0285)
Hispanic	-0.155***	-0.0369	-0.148***	-0.0370	-0.0341	0.0962***	-0.0250	0.0938***
	(0.0298)	(0.0291)	(0.0299)	(0.0291)	(0.0310)	(0.0304)	(0.0307)	(0.0303)
Age	0.0305***	0.0113	0.0270**	0.00940	0.00244	0.000564	-0.00533	-0.00422
	(0.0115)	(0.0110)	(0.0115)	(0.0110)	(0.0127)	(0.0118)	(0.0126)	(0.0118)
Arith Reason		0.189***		0.185***		0.216***		0.203***
		(0.0121)		(0.0123)		(0.0147)		(0.0147)
Arith Reason^2		-0.00296		-0.00224		-0.00349		-0.00285
		(0.0107)		(0.0108)		(0.0139)		(0.0138)
Rotter			-0.0522***	-0.0286***			-0.0812***	-0.0505***
			(0.0113)	(0.0103)			(0.0136)	(0.0126)
Rotter^2			0.00423	0.0101			0.00416	0.00511
			(0.00777)	(0.00713)			(0.00880)	(0.00817)
Observations	13163	12981	13137	12955	11819	11640	11765	11586
R^2	0.111	0.200	0.119	0.203	0.048	0.131	0.065	0.136

	Men (1-4)				Women (5-8)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	-0.322***	-0.143***	-0.331***	-0.166***	-0.180***	0.0105	-0.207***	-0.0286
	(0.0241)	(0.0252)	(0.0235)	(0.0253)	(0.0266)	(0.0285)	(0.0261)	(0.0286)
Hispanic	-0.155***	-0.0369	-0.139***	-0.0387	-0.0341	0.0962***	-0.0281	0.0868***
	(0.0298)	(0.0291)	(0.0287)	(0.0286)	(0.0310)	(0.0304)	(0.0295)	(0.0295)
Age	0.0305***	0.0113	0.0149	0.00402	0.00244	0.000564	-0.00517	-0.00448
	(0.0115)	(0.0110)	(0.0111)	(0.0107)	(0.0127)	(0.0118)	(0.0121)	(0.0115)
Arith Reason		0.189***		0.170***		0.216***		0.193***
		(0.0121)		(0.0124)		(0.0147)		(0.0147)
Arith Reason^2		-0.00296		0.000900		-0.00349		-0.00187
		(0.0107)		(0.0107)		(0.0139)		(0.0135)
Rosenberg			0.111***	0.0654***			0.114***	0.0807***
			(0.0111)	(0.0111)			(0.0125)	(0.0121)
Rosenberg^2			-0.00928	-0.00705			-0.0177	-0.0103
			(0.0104)	(0.00996)			(0.0121)	(0.0115)
Observations	13163	12981	13153	12971	11819	11640	11809	11630
R^2	0.111	0.200	0.144	0.211	0.048	0.131	0.084	0.148

Regressions include annual time dummy variables. Standard errors in parentheses correct for the longitudinal structure of the NLSY by accounting for repeated observations of individuals over time. *** p<0.01, ** p<0.05, * p<0.1

Table 2.11 Log Wage Regression Using Pooled Data with Paragraph Comprehension, Locus of Control, and Self-Esteem

	Men (1-4)				Women (5-8)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	-0.322***	-0.163***	-0.317***	-0.167***	-0.180***	0.0101	-0.157***	0.0146
	(0.0241)	(0.0251)	(0.0243)	(0.0252)	(0.0266)	(0.0268)	(0.0265)	(0.0267)
Hispanic	-0.155***	-0.0444	-0.148***	-0.0442	-0.0341	0.100***	-0.0250	0.0986***
	(0.0298)	(0.0295)	(0.0299)	(0.0295)	(0.0310)	(0.0295)	(0.0307)	(0.0293)
Age	0.0305***	0.0107	0.0270**	0.00843	0.00244	-0.00330	-0.00533	-0.00811
	(0.0115)	(0.0108)	(0.0115)	(0.0108)	(0.0127)	(0.0116)	(0.0126)	(0.0115)
Parag Comp		0.186***		0.182***		0.225***		0.213***
		(0.0138)		(0.0139)		(0.0147)		(0.0147)
Parag Comp ^{2}		0.0134		0.0125		0.0191		0.0151
		(0.00987)		(0.00989)		(0.0123)		(0.0124)
Rotter			-0.0522***	-0.0318***			-0.0812***	-0.0523***
			(0.0113)	(0.0106)			(0.0136)	(0.0128)
Rotter^2			0.00423	0.0127*			0.00416	0.00409
			(0.00777)	(0.00718)			(0.00880)	(0.00836)
Observations	13163	12981	13137	12955	11819	11640	11765	11586
R^2	0.111	0.197	0.119	0.200	0.048	0.146	0.065	0.151

	Men (1-4)				Women (5-8)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	-0.322***	-0.163***	-0.331***	-0.185***	-0.180***	0.0101	-0.207***	-0.0251
	(0.0241)	(0.0251)	(0.0235)	(0.0252)	(0.0266)	(0.0268)	(0.0261)	(0.0272)
Hispanic	-0.155***	-0.0444	-0.139***	-0.0470	-0.0341	0.100***	-0.0281	0.0910***
	(0.0298)	(0.0295)	(0.0287)	(0.0292)	(0.0310)	(0.0295)	(0.0295)	(0.0289)
Age	0.0305***	0.0107	0.0149	0.00418	0.00244	-0.00330	-0.00517	-0.00734
	(0.0115)	(0.0108)	(0.0111)	(0.0107)	(0.0127)	(0.0116)	(0.0121)	(0.0113)
Parag Comp		0.186***		0.169***		0.225***		0.204***
		(0.0138)		(0.0143)		(0.0147)		(0.0148)
Parag Comp ^{2}		0.0134		0.0139		0.0191		0.0209*
		(0.00987)		(0.00986)		(0.0123)		(0.0122)
Rosenberg			0.111***	0.0607***			0.114***	0.0708***
			(0.0111)	(0.0112)			(0.0125)	(0.0123)
Rosenberg^2			-0.00928	-0.00650			-0.0177	-0.00913
			(0.0104)	(0.0100)			(0.0121)	(0.0115)
Observations	13163	12981	13153	12971	11819	11640	11809	11630
R^2	0.111	0.197	0.144	0.206	0.048	0.146	0.084	0.159

Regressions include annual time dummy variables. Standard errors in parentheses correct for the longitudinal structure of the NLSY by accounting for repeated observations of individuals over time. *** p<0.01, ** p<0.05, * p<0.1

Table 2.12 Log Wage Regression Using Pooled Data with Numerical Operations, Locus of Control, and Self-Esteem

	Men (1-4)				Women (5-8)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	-0.322***	-0.188***	-0.317***	-0.190***	-0.180***	-0.0577**	-0.157***	-0.0468*
	(0.0241)	(0.0246)	(0.0243)	(0.0247)	(0.0266)	(0.0267)	(0.0265)	(0.0267)
Hispanic	-0.155***	-0.0689**	-0.148***	-0.0672**	-0.0341	0.0388	-0.0250	0.0409
	(0.0298)	(0.0285)	(0.0299)	(0.0285)	(0.0310)	(0.0297)	(0.0307)	(0.0295)
Age	0.0305***	0.00944	0.0270**	0.00753	0.00244	-0.00316	-0.00533	-0.00943
	(0.0115)	(0.0108)	(0.0115)	(0.0108)	(0.0127)	(0.0120)	(0.0126)	(0.0119)
Num Oper		0.181***		0.176***		0.176***		0.168***
		(0.0123)		(0.0123)		(0.0125)		(0.0124)
Num Oper^2		-0.0106		-0.0113		-0.0249**		-0.0247**
		(0.00902)		(0.00902)		(0.0107)		(0.0106)
Rotter			-0.0522***	-0.0321***			-0.0812***	-0.0637***
			(0.0113)	(0.0105)			(0.0136)	(0.0130)
Rotter^2			0.00423	0.00606			0.00416	0.00152
			(0.00777)	(0.00728)			(0.00880)	(0.00831)
Observations	13163	12981	13137	12955	11819	11640	11765	11586
R^2	0.111	0.200	0.119	0.203	0.048	0.120	0.065	0.129

	Men (1-4)				Women (5-8)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	-0.322***	-0.188***	-0.331***	-0.207***	-0.180***	-0.0577**	-0.207***	-0.0913***
	(0.0241)	(0.0246)	(0.0235)	(0.0246)	(0.0266)	(0.0267)	(0.0261)	(0.0266)
Hispanic	-0.155***	-0.0689**	-0.139***	-0.0684**	-0.0341	0.0388	-0.0281	0.0358
	(0.0298)	(0.0285)	(0.0287)	(0.0281)	(0.0310)	(0.0297)	(0.0295)	(0.0287)
Age	0.0305***	0.00944	0.0149	0.00248	0.00244	-0.00316	-0.00517	-0.00819
	(0.0115)	(0.0108)	(0.0111)	(0.0107)	(0.0127)	(0.0120)	(0.0121)	(0.0116)
Num Oper		0.181***		0.164***		0.176***		0.157***
		(0.0123)		(0.0127)		(0.0125)		(0.0124)
Num Oper^2		-0.0106		-0.00867		-0.0249**		-0.0219**
		(0.00902)		(0.00898)		(0.0107)		(0.0106)
Rosenberg			0.111***	0.0635***			0.114***	0.0866***
			(0.0111)	(0.0111)			(0.0125)	(0.0121)
Rosenberg^2			-0.00928	-0.00534			-0.0177	-0.0141
			(0.0104)	(0.00995)			(0.0121)	(0.0116)
Observations	13163	12981	13153	12971	11819	11640	11809	11630
R^2	0.111	0.200	0.144	0.211	0.048	0.120	0.084	0.140

Regressions include annual time dummy variables. Standard errors in parentheses correct for the longitudinal structure of the NLSY by accounting for repeated observations of individuals over time. *** p<0.01, ** p<0.05, * p<0.1

Table 2.13 Log Wage Regression Using Pooled Data with Locus of Control and Race Interactions

	Men (1-4)				Women (5-8)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	-0.322***	-0.0806**	-0.306***	-0.0732**	-0.180***	0.0602*	-0.170***	0.0521
	(0.0241)	(0.0317)	(0.0300)	(0.0348)	(0.0266)	(0.0310)	(0.0325)	(0.0350)
Hispanic	-0.155***	0.00592	-0.177***	-0.0162	-0.0341	0.0974***	-0.0278	0.100**
	(0.0298)	(0.0359)	(0.0374)	(0.0415)	(0.0310)	(0.0363)	(0.0372)	(0.0404)
Age	0.0305***	0.00259	0.0274**	0.00173	0.00244	-0.00688	-0.00496	-0.00999
	(0.0115)	(0.0106)	(0.0115)	(0.0107)	(0.0127)	(0.0112)	(0.0126)	(0.0112)
AFQT		0.248***		0.243***		0.258***		0.248***
		(0.0168)		(0.0172)		(0.0207)		(0.0207)
AFQT x Black		0.0529		0.0499		0.111***		0.0996**
		(0.0486)		(0.0487)		(0.0413)		(0.0425)
AFQT x Hispanic		-0.0210		-0.0151		0.0718*		0.0732*
		(0.0421)		(0.0426)		(0.0412)		(0.0411)
$AFQT^2$		0.0836***		0.0799***		0.0659***		0.0622***
		(0.0167)		(0.0171)		(0.0167)		(0.0162)
$AFQT^2$ x Black		-0.0272		-0.0269		0.00659		0.00677
		(0.0434)		(0.0431)		(0.0350)		(0.0358)
$AFQT^2$ x Hispanic		-0.0500		-0.0442		0.0560		0.0576
		(0.0415)		(0.0413)		(0.0477)		(0.0477)
Rotter			-0.0667***	-0.0255*			-0.0778***	-0.0414**
			(0.0160)	(0.0138)			(0.0191)	(0.0174)
Rotter x Black			0.0235	0.00643			-0.0272	-0.00475
			(0.0238)	(0.0212)			(0.0328)	(0.0311)
Rotter x Hispanic			0.0327	0.00936			0.0137	0.00450
			(0.0375)	(0.0347)			(0.0339)	(0.0302)
$Rotter^2$			0.00302	0.00807			-0.00151	0.00216
			(0.0120)	(0.0101)			(0.0132)	(0.0119)
$Rotter^2$ x Black			-0.0106	-0.0111			0.0213	0.00794
			(0.0163)	(0.0147)			(0.0211)	(0.0192)
$Rotter^2$ x Hispanic			0.0222	0.0160			2.42e-05	-0.00583
			(0.0264)	(0.0245)			(0.0216)	(0.0200)
Observations	13163	13163	13137	13137	11819	11819	11765	11765
R^2	0.111	0.224	0.121	0.226	0.048	0.166	0.066	0.168

Regressions include annual time dummy variables. Standard errors in parentheses correct for the longitudinal structure of the NLSY by accounting for repeated observations of individuals over time. *** p<0.01, ** p<0.05, * p<0.1

Table 2.14 Log Wage Regression Using Pooled Data with Self-Esteem and Race Interactions

	Men (1-4)				Women (5-8)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	-0.322***	-0.0806**	-0.335***	-0.103***	-0.180***	0.0602*	-0.226***	0.0208
	(0.0241)	(0.0317)	(0.0320)	(0.0374)	(0.0266)	(0.0310)	(0.0363)	(0.0395)
Hispanic	-0.155***	0.00592	-0.0859**	0.0580	-0.0341	0.0974***	-0.0366	0.0845*
	(0.0298)	(0.0359)	(0.0406)	(0.0425)	(0.0310)	(0.0363)	(0.0406)	(0.0437)
Age	0.0305***	0.00259	0.0152	-0.00231	0.00244	-0.00688	-0.00715	-0.0118
	(0.0115)	(0.0106)	(0.0111)	(0.0105)	(0.0127)	(0.0112)	(0.0121)	(0.0109)
AFQT		0.248***		0.238***		0.258***		0.246***
		(0.0168)		(0.0175)		(0.0207)		(0.0213)
AFQT x Black		0.0529		0.0360		0.111***		0.0902**
		(0.0486)		(0.0504)		(0.0413)		(0.0419)
AFQT x Hispanic		-0.0210		-0.0698		0.0718*		0.0259
		(0.0421)		(0.0461)		(0.0412)		(0.0412)
$AFQT^2$		0.0836***		0.0787***		0.0659***		0.0640***
		(0.0167)		(0.0167)		(0.0167)		(0.0168)
$AFQT^2$ x Black		-0.0272		-0.0299		0.00659		0.0101
		(0.0434)		(0.0422)		(0.0350)		(0.0343)
$AFQT^2$ x Hispanic		-0.0500		-0.0645*		0.0560		0.0488
		(0.0415)		(0.0391)		(0.0477)		(0.0444)
Rosenberg			0.0989***	0.0421***			0.0739***	0.0246
			(0.0150)	(0.0144)			(0.0183)	(0.0171)
Rosenberg x Black			0.00713	-0.00165			0.0719***	0.0624**
			(0.0252)	(0.0251)			(0.0277)	(0.0264)
Rosenberg x Hispanic			0.0441	0.0498			0.0957***	0.0921***
			(0.0318)	(0.0337)			(0.0323)	(0.0309)
$Rosenberg^2$			-0.000698	0.00273			-0.0364**	-0.0272*
			(0.0147)	(0.0134)			(0.0175)	(0.0161)
$Rosenberg^2$ x Black			0.00643	0.00379			0.0327	0.0154
			(0.0231)	(0.0222)			(0.0280)	(0.0261)
$Rosenberg^2$ x Hispanic			-0.0451	-0.0434			0.0337	0.0311
			(0.0288)	(0.0273)			(0.0294)	(0.0272)
Observations	13163	13163	13153	13153	11819	11819	11809	11809
R^2	0.111	0.224	0.146	0.233	0.048	0.166	0.090	0.179

Regressions include annual time dummy variables. Standard errors in parentheses correct for the longitudinal structure of the NLSY by accounting for repeated observations of individuals over time. *** p<0.01, ** p<0.05, * p<0.1

Table 2.15 Log Wage Regression Using Pooled Data with Locus of Control, Self-Esteem, and South

	Men (1-4)				Women (5-8)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	-0.292***	-0.107***	-0.287***	-0.111***	-0.145***	0.0724***	-0.122***	0.0744***
	(0.0257)	(0.0256)	(0.0258)	(0.0257)	(0.0282)	(0.0273)	(0.0278)	(0.0272)
Hispanic	-0.158***	-0.0199	-0.151***	-0.0201	-0.0315	0.133***	-0.0222	0.131***
	(0.0298)	(0.0286)	(0.0298)	(0.0286)	(0.0308)	(0.0291)	(0.0306)	(0.0289)
Age	0.0303***	0.00311	0.0268**	0.00168	0.00230	-0.00495	-0.00541	-0.00878
	(0.0114)	(0.0107)	(0.0114)	(0.0107)	(0.0126)	(0.0111)	(0.0125)	(0.0111)
South	-0.0892***	-0.0420**	-0.0888***	-0.0424**	-0.105***	-0.0756***	-0.104***	-0.0765***
	(0.0236)	(0.0208)	(0.0234)	(0.0208)	(0.0258)	(0.0229)	(0.0252)	(0.0227)
AFQT		0.256***		0.252***		0.290***		0.277***
		(0.0139)		(0.0141)		(0.0162)		(0.0162)
$AFQT^2$		0.0641***		0.0613***		0.0639***		0.0618***
		(0.0137)		(0.0137)		(0.0146)		(0.0145)
Rotter			-0.0518***	-0.0225**			-0.0814***	-0.0432***
			(0.0113)	(0.0100)			(0.0135)	(0.0123)
$Rotter^2$			0.00349	0.00697			0.00486	0.00408
			(0.00773)	(0.00693)			(0.00875)	(0.00793)
Observations	13163	13163	13137	13137	11819	11819	11765	11765
R^2	0.116	0.223	0.124	0.225	0.055	0.166	0.072	0.169

	Men (1-4)				Women (5-8)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	-0.292***	-0.107***	-0.304***	-0.127***	-0.145***	0.0724***	-0.176***	0.0375
	(0.0257)	(0.0256)	(0.0250)	(0.0258)	(0.0282)	(0.0273)	(0.0279)	(0.0280)
Hispanic	-0.158***	-0.0199	-0.142***	-0.0232	-0.0315	0.133***	-0.0260	0.123***
	(0.0298)	(0.0286)	(0.0287)	(0.0283)	(0.0308)	(0.0291)	(0.0295)	(0.0285)
Age	0.0303***	0.00311	0.0149	-0.00224	0.00230	-0.00495	-0.00510	-0.00834
	(0.0114)	(0.0107)	(0.0111)	(0.0106)	(0.0126)	(0.0111)	(0.0120)	(0.0108)
South	-0.0892***	-0.0420**	-0.0794***	-0.0413**	-0.105***	-0.0756***	-0.0883***	-0.0691***
	(0.0236)	(0.0208)	(0.0228)	(0.0207)	(0.0258)	(0.0229)	(0.0252)	(0.0227)
AFQT		0.256***		0.237***		0.290***		0.266***
		(0.0139)		(0.0147)		(0.0162)		(0.0166)
$AFQT^2$		0.0641***		0.0598***		0.0639***		0.0602***
		(0.0137)		(0.0133)		(0.0146)		(0.0144)
Rosenberg			0.109***	0.0525***			0.111***	0.0594***
			(0.0110)	(0.0110)			(0.0125)	(0.0118)
$Rosenberg^2$			-0.00836	-0.00478			-0.0155	-0.0101
			(0.0103)	(0.00971)			(0.0120)	(0.0111)
Observations	13163	13163	13153	13153	11819	11819	11809	11809
R^2	0.116	0.223	0.148	0.231	0.055	0.166	0.089	0.175

Regressions include annual time dummy variables. Standard errors in parentheses correct for the longitudinal structure of the NLSY by accounting for repeated observations of individuals over time.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 2.16 Log Wage Regression Using Pooled Data with Locus of Control and South Interactions

	Men (1-4)				Women (5-8)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	-0.292***	-0.116***	-0.303***	-0.115***	-0.145***	0.0576**	-0.146***	0.0650**
	(0.0257)	(0.0254)	(0.0249)	(0.0256)	(0.0282)	(0.0273)	(0.0275)	(0.0274)
Hispanic	-0.158***	-0.0189	-0.148***	-0.0195	-0.0315	0.126***	-0.0246	0.125***
	(0.0298)	(0.0289)	(0.0299)	(0.0288)	(0.0308)	(0.0290)	(0.0306)	(0.0289)
Age	0.0303***	0.00267	0.0258**	0.000826	0.00230	-0.00515	-0.00639	-0.00949
	(0.0114)	(0.0107)	(0.0115)	(0.0107)	(0.0126)	(0.0111)	(0.0126)	(0.0111)
South	-0.0892***				-0.105***			
	(0.0236)				(0.0258)			
AFQT		0.253***		0.249***		0.258***		0.249***
		(0.0165)		(0.0167)		(0.0205)		(0.0206)
AFQT x South		0.0204		0.0183		0.0947***		0.0859***
		(0.0266)		(0.0271)		(0.0296)		(0.0304)
$AFQT^2$		0.0661***		0.0602***		0.0692***		0.0634***
		(0.0148)		(0.0154)		(0.0174)		(0.0177)
$AFQT^2$ x South		-0.00214		0.00832		-0.00440		0.00762
		(0.0236)		(0.0252)		(0.0236)		(0.0244)
Rotter			-0.0532***	-0.0230*			-0.0663***	-0.0335**
			(0.0146)	(0.0129)			(0.0179)	(0.0167)
Rotter x South			0.00530	0.00324			-0.0366	-0.0194
			(0.0224)	(0.0203)			(0.0269)	(0.0246)
$Rotter^2$			0.0165*	0.0144			0.0102	0.00928
			(0.00959)	(0.00881)			(0.0111)	(0.0105)
$Rotter^2$ x South			-0.0350***	-0.0197			-0.0158	-0.0165
			(0.0132)	(0.0127)			(0.0155)	(0.0146)
Observations	13163	13163	13137	13137	11819	11819	11765	11765
R^2	0.116	0.222	0.122	0.224	0.055	0.166	0.067	0.169

Regressions include annual time dummy variables. Standard errors in parentheses correct for the longitudinal structure of the NLSY by accounting for repeated observations of individuals over time. *** p<0.01, ** p<0.05, * p<0.1

Table 2.17 Log Wage Regression Using Pooled Data with Self-Esteem and South Interactions

	Men (1-4)				Women (5-8)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Black	-0.292***	-0.116***	-0.318***	-0.134***	-0.145***	0.0576**	-0.190***	0.0354
	(0.0257)	(0.0254)	(0.0242)	(0.0256)	(0.0282)	(0.0273)	(0.0265)	(0.0277)
Hispanic	-0.158***	-0.0189	-0.141***	-0.0225	-0.0315	0.126***	-0.0245	0.119***
	(0.0298)	(0.0289)	(0.0287)	(0.0285)	(0.0308)	(0.0290)	(0.0294)	(0.0282)
Age	0.0303***	0.00267	0.0150	-0.00237	0.00230	-0.00515	-0.00467	-0.00812
	(0.0114)	(0.0107)	(0.0111)	(0.0106)	(0.0126)	(0.0111)	(0.0120)	(0.0108)
South	-0.0892***				-0.105***			
	(0.0236)				(0.0258)			
AFQT		0.253***		0.235***		0.258***		0.232***
		(0.0165)		(0.0173)		(0.0205)		(0.0208)
AFQT x South		0.0204		0.0166		0.0947***		0.102***
		(0.0266)		(0.0275)		(0.0296)		(0.0298)
$AFQT^2$		0.0661***		0.0573***		0.0692***		0.0525***
		(0.0148)		(0.0150)		(0.0174)		(0.0181)
$AFQT^2$ x South		-0.00214		0.0107		-0.00440		0.0311
		(0.0236)		(0.0261)		(0.0236)		(0.0247)
Rosenberg			0.108***	0.0557***			0.1000***	0.0559***
			(0.0138)	(0.0136)			(0.0160)	(0.0153)
Rosenberg x South			0.00146	-0.0102			0.0309	0.00608
			(0.0226)	(0.0226)			(0.0253)	(0.0231)
$Rosenberg^2$			0.00440	0.00177			0.00566	0.0145
			(0.0119)	(0.0114)			(0.0144)	(0.0140)
$Rosenberg^2$ x South			-0.0370**	-0.0195			-0.0521***	-0.0616***
			(0.0167)	(0.0181)			(0.0185)	(0.0178)
Observations	13163	13163	13153	13153	11819	11819	11809	11809
R^2	0.116	0.222	0.146	0.230	0.055	0.166	0.089	0.180

Regressions include annual time dummy variables. Standard errors in parentheses correct for the longitudinal structure of the NLSY by accounting for repeated observations of individuals over time. *** p<0.01, ** p<0.05, * p<0.1

Figure 2.1 Kernel Density Estimation of AFQT, Rotter, and Rosenberg Scores

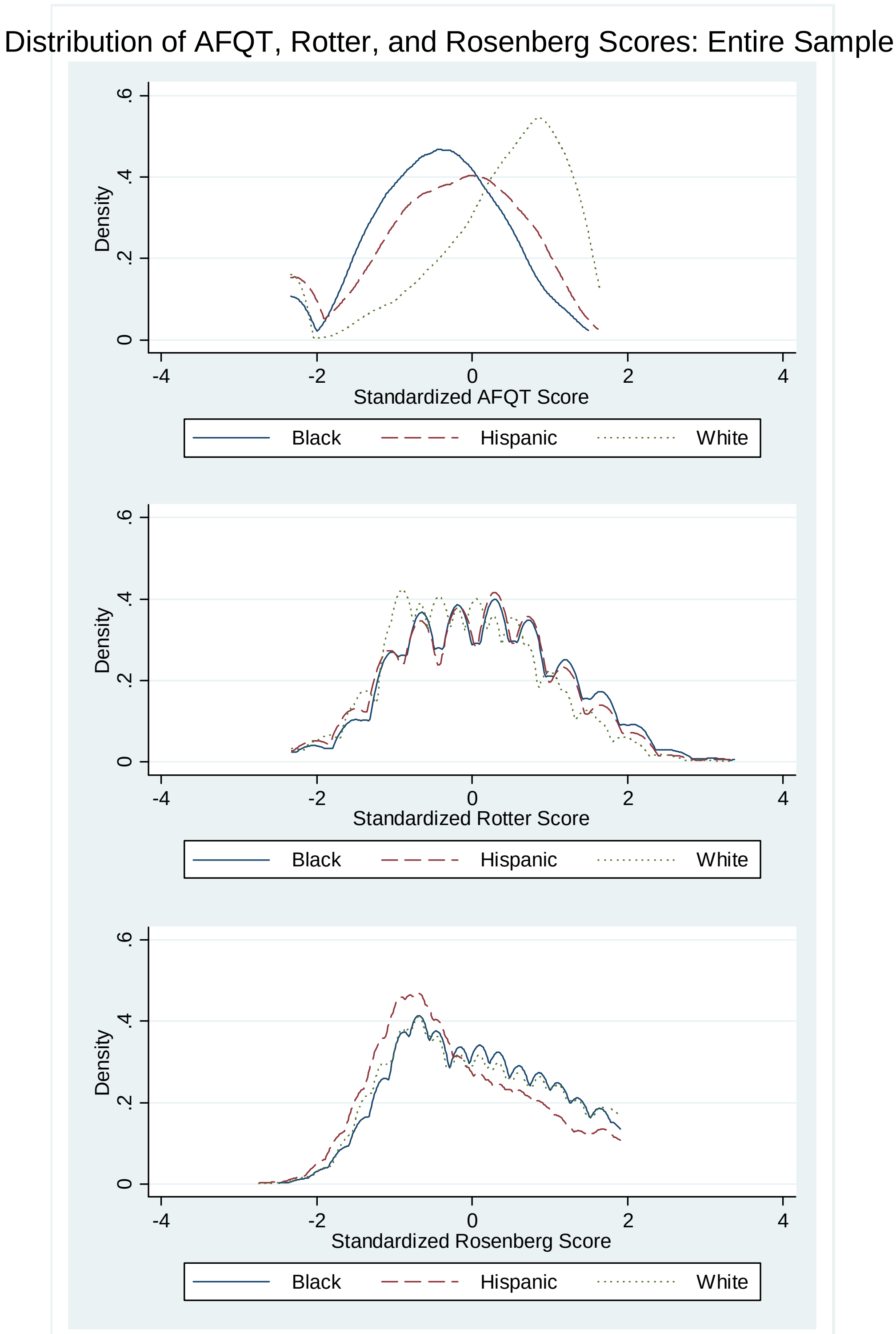


Figure 2.1 (continued)

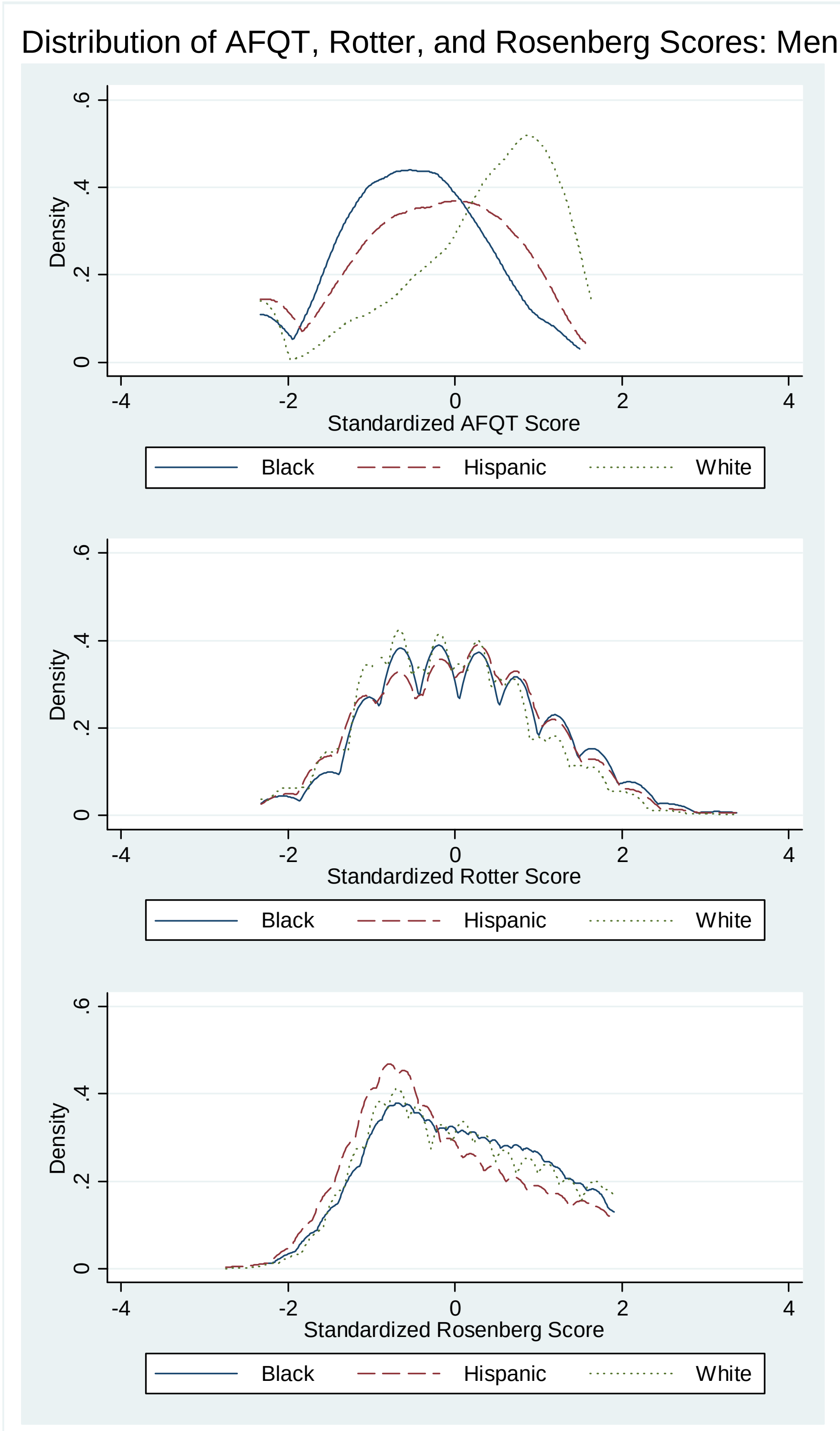


Figure 2.1 (continued)

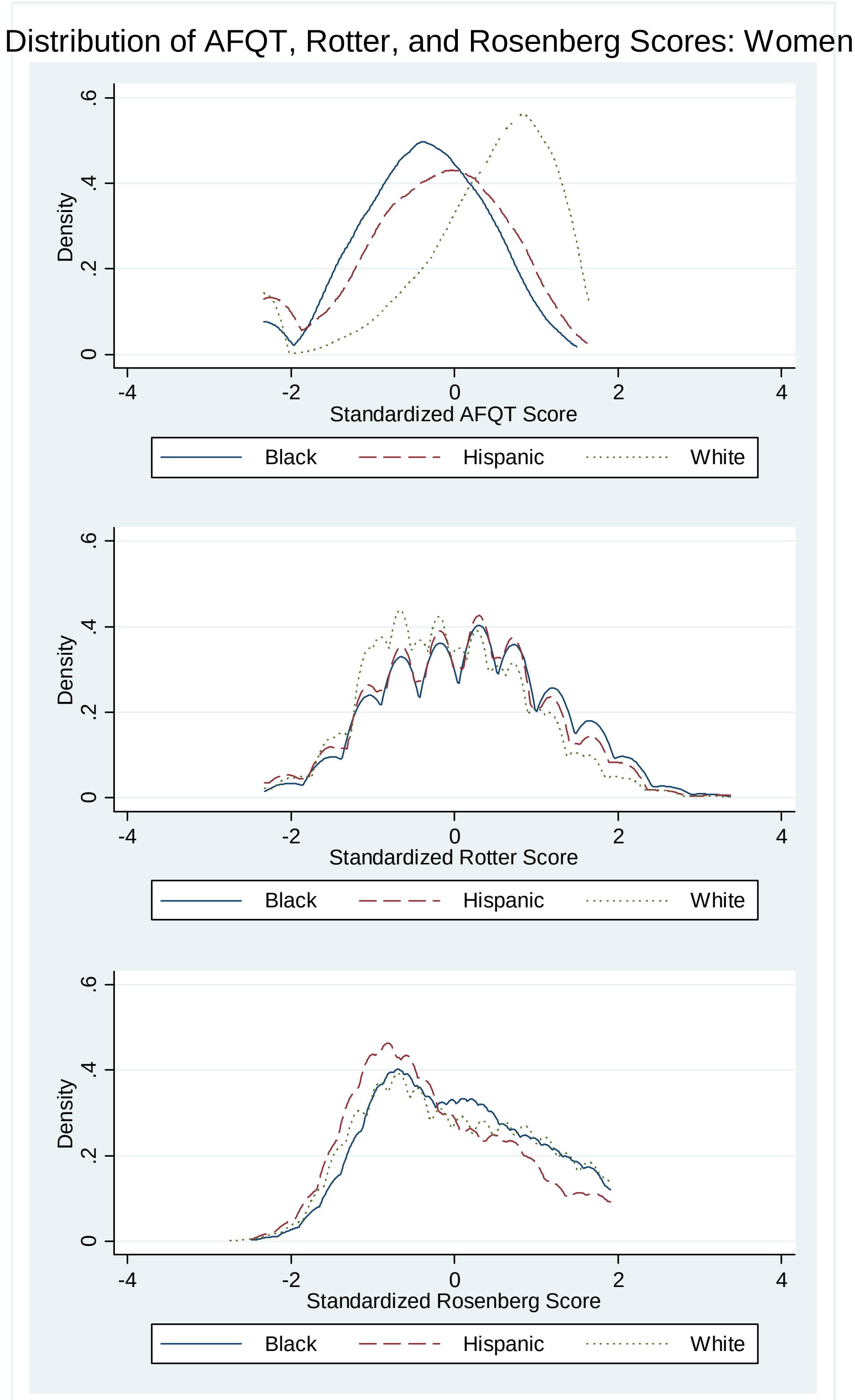


Figure 2.2 Change in Wage Gap, Pooled Data

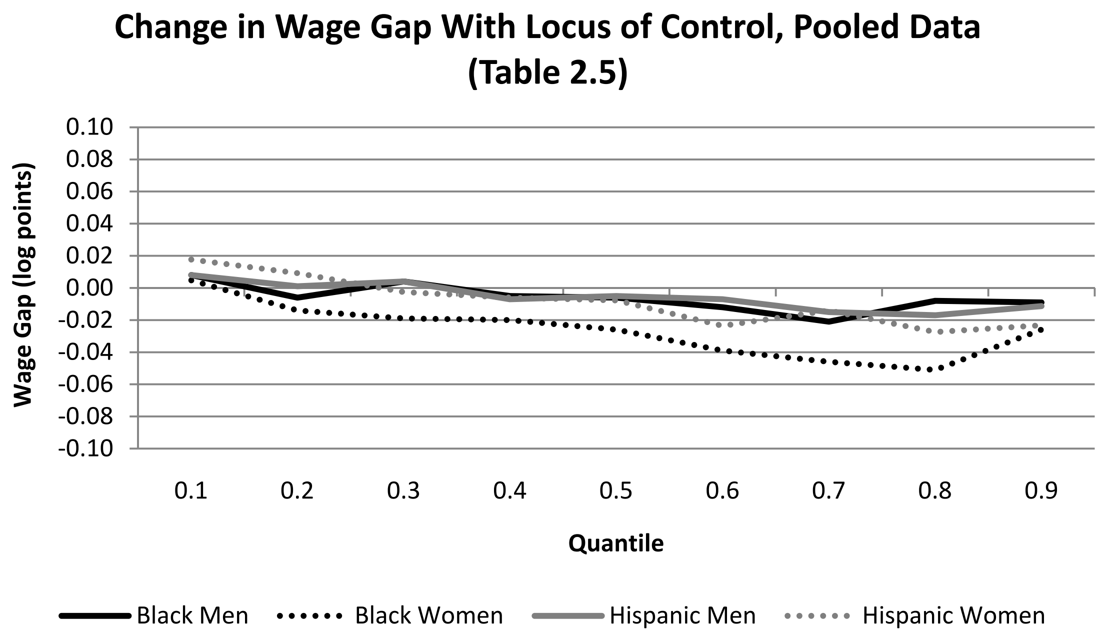
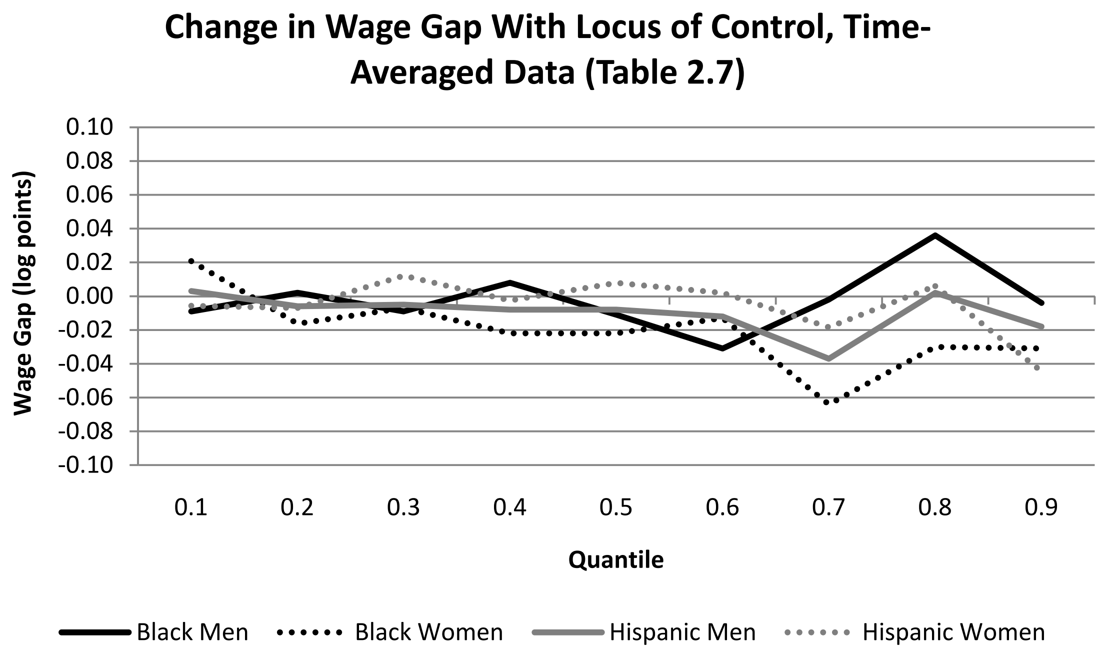


Figure 2.3 Change in Wage Gap, Time-Averaged Data



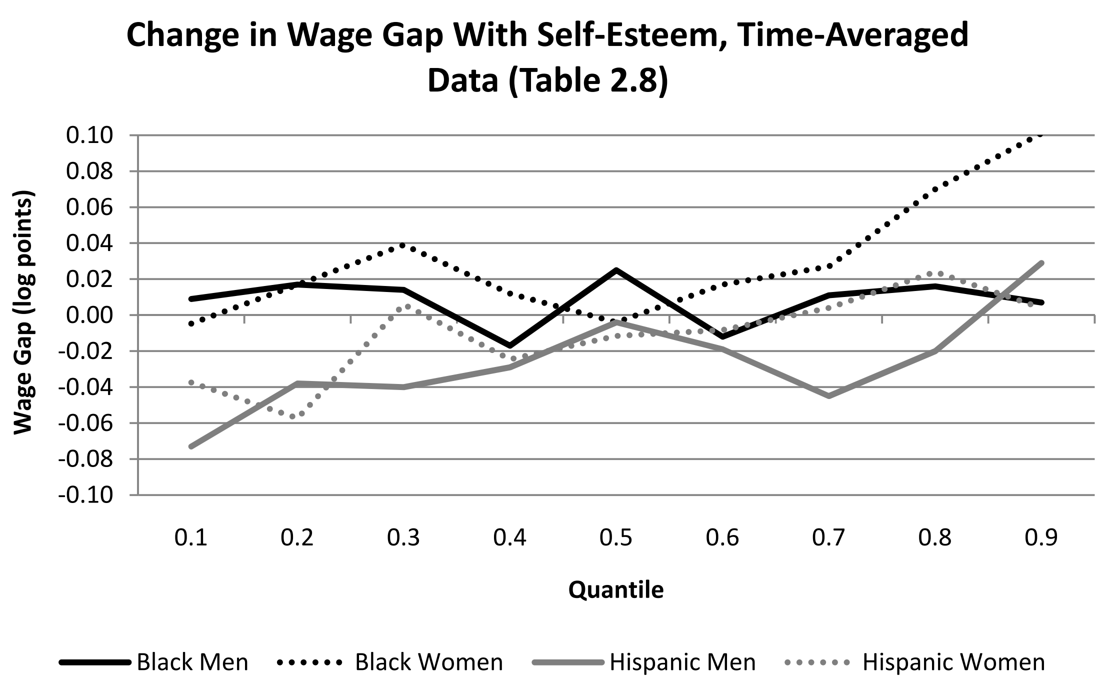
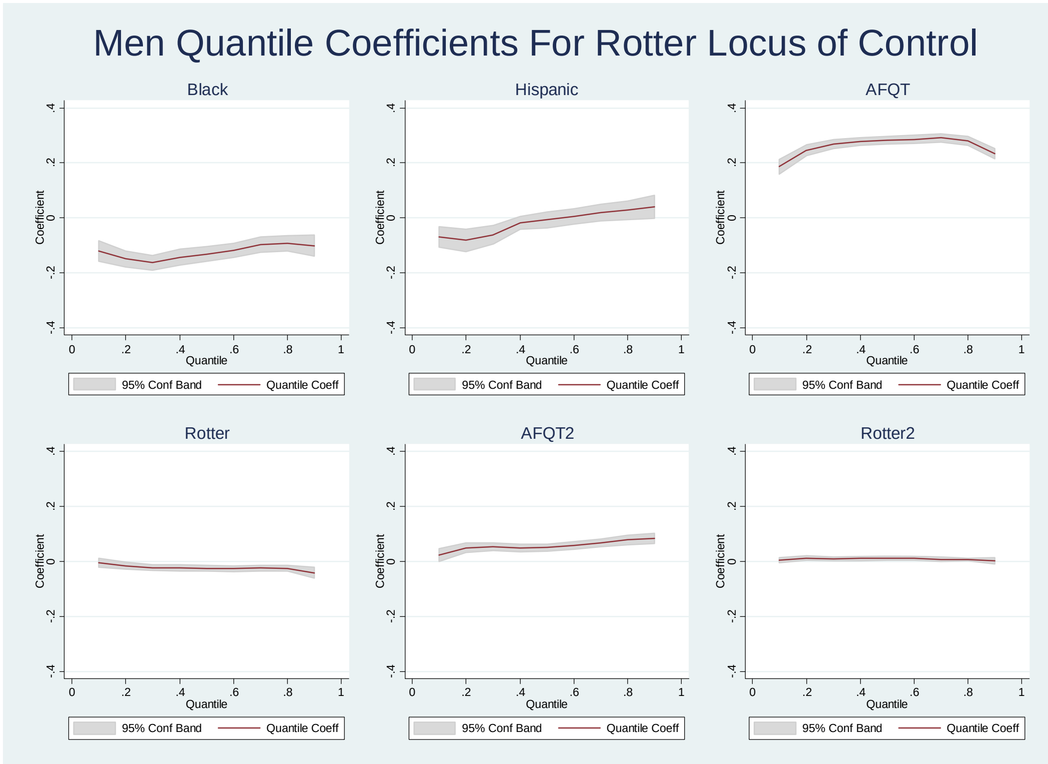


Figure 2.4 Coefficients From Quantile Wage Regression, Pooled Data



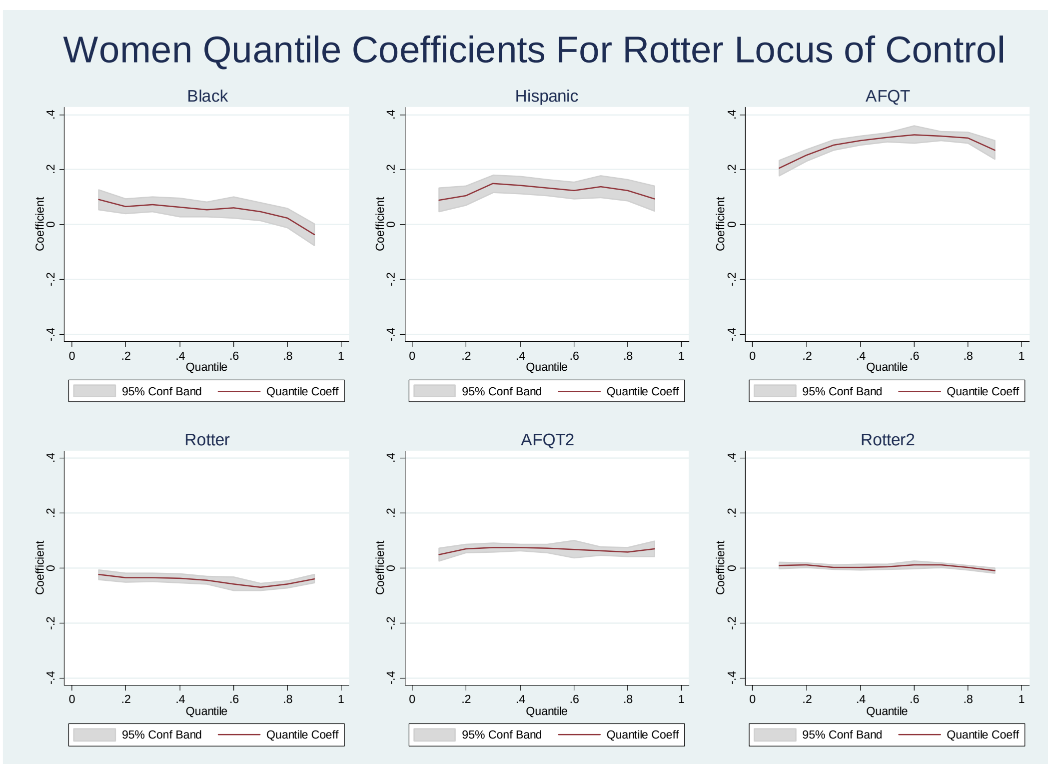
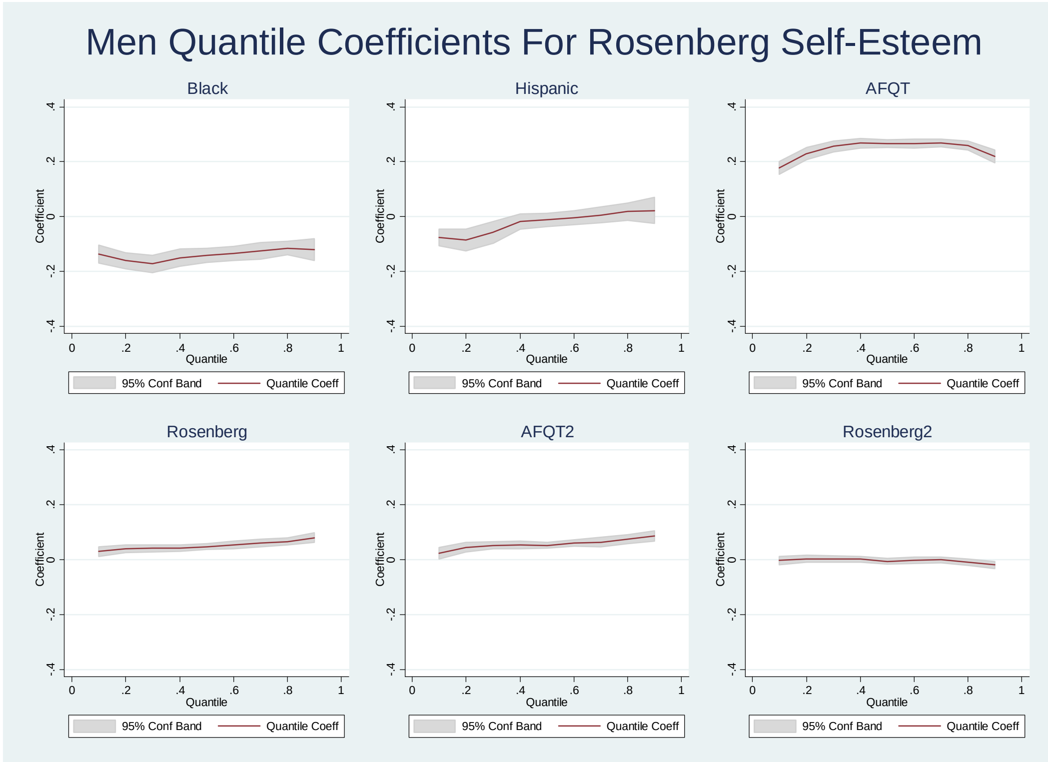


Figure 2.4 (continued)



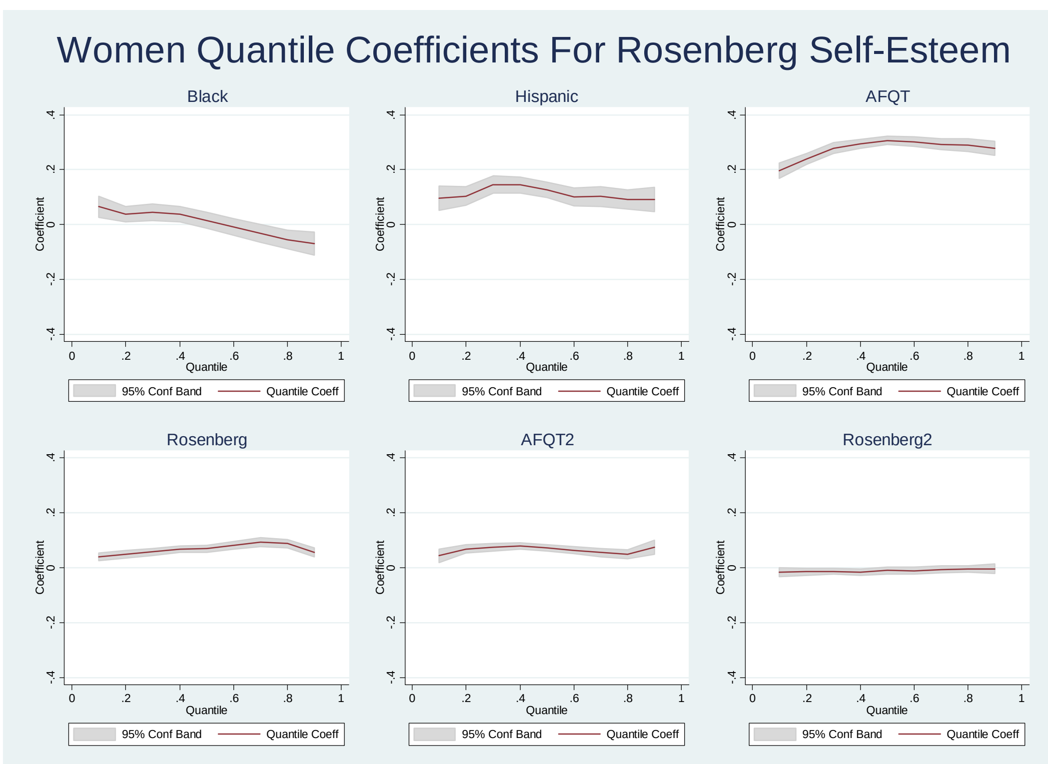
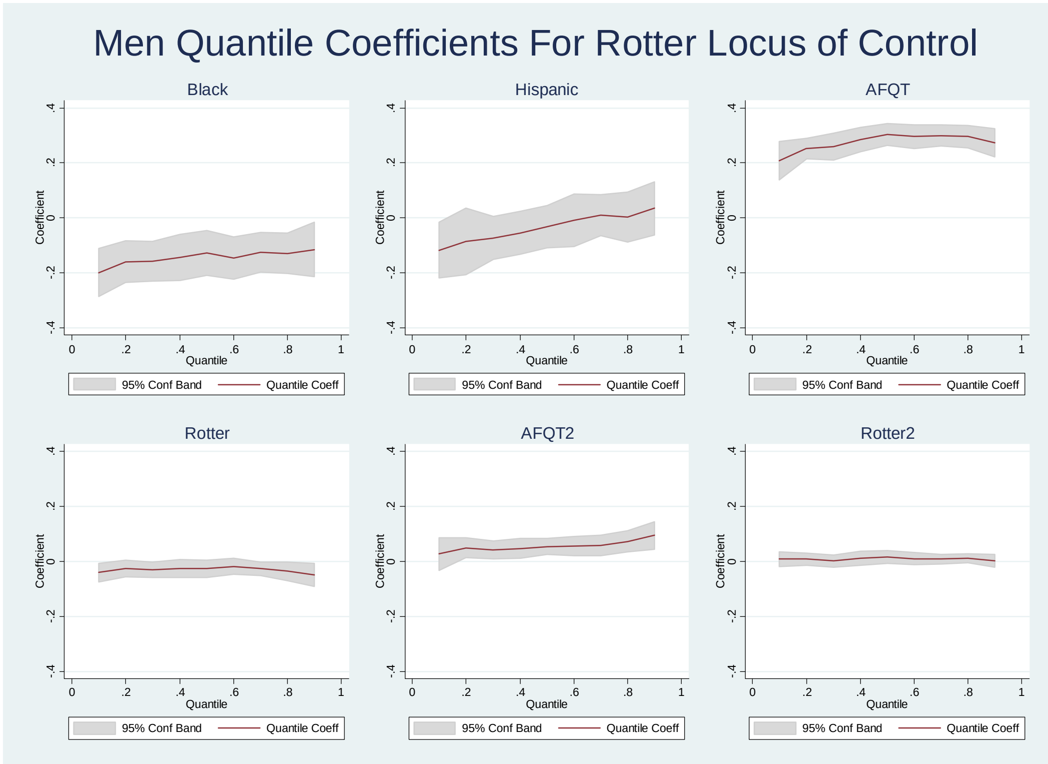


Figure 2.5 Coefficients From Quantile Wage Regression, Time-Averaged Data



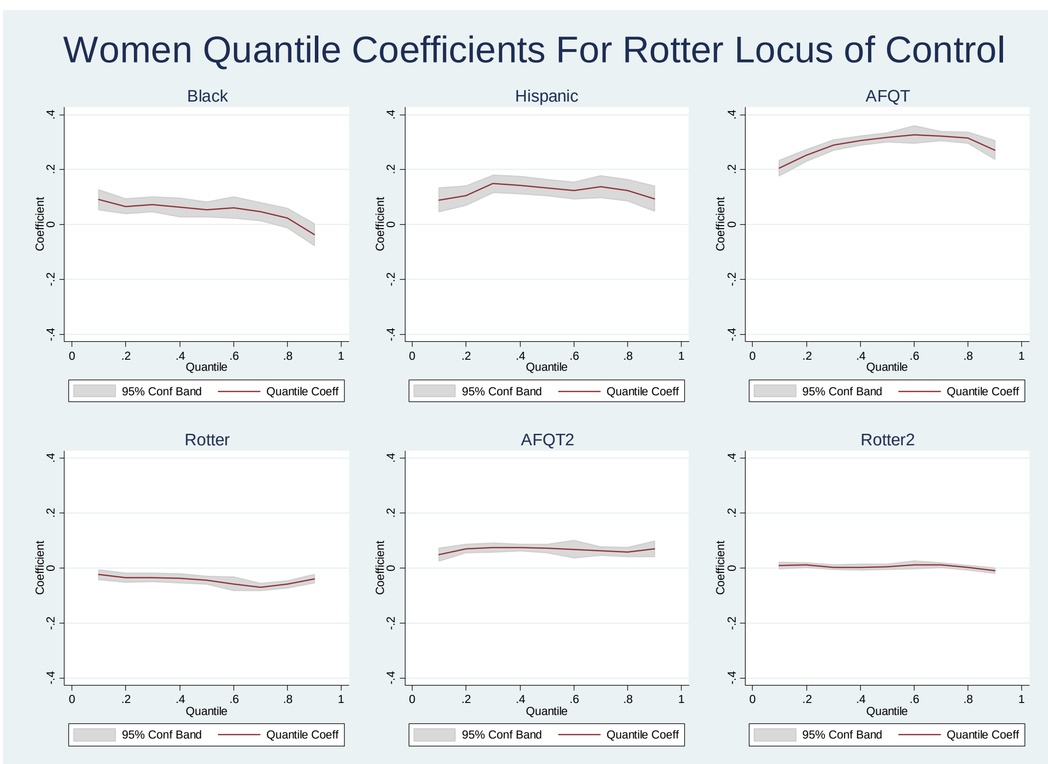
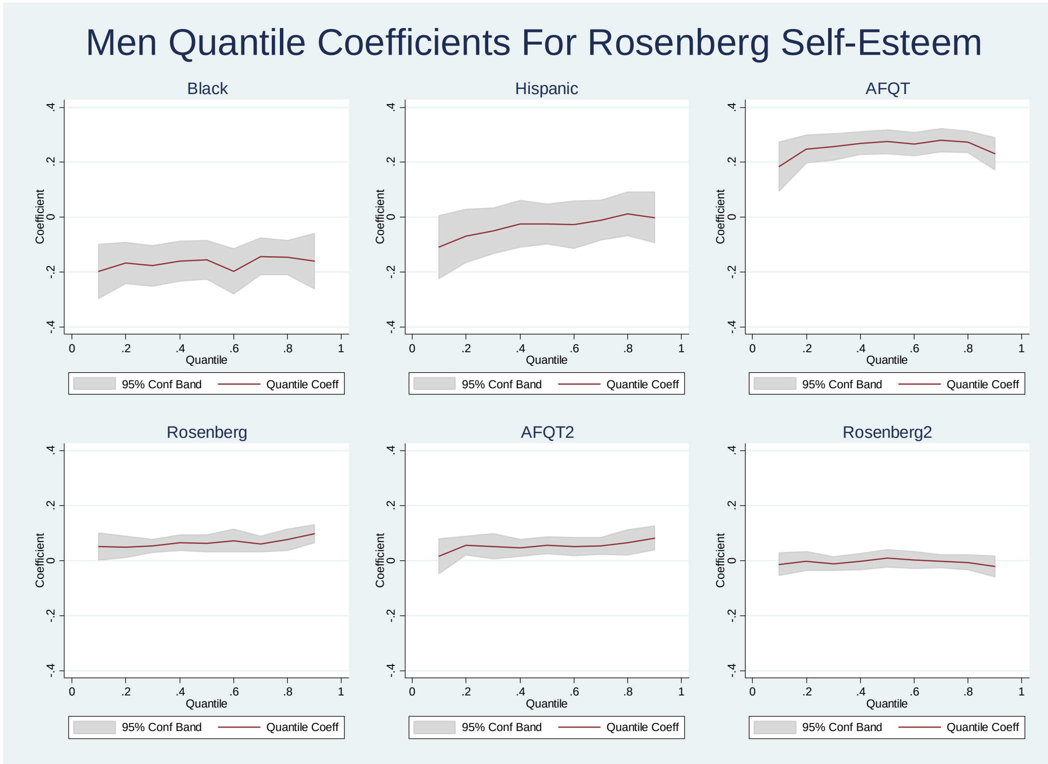
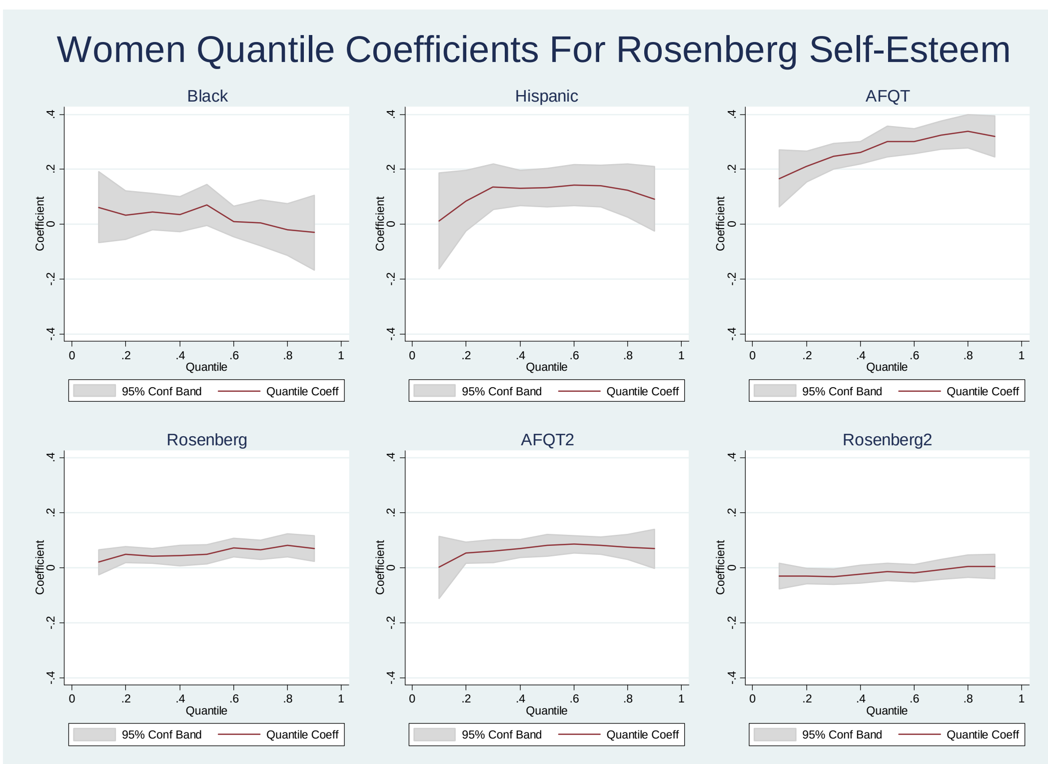


Figure 2.5 (continued)





3 AN EXPLORATION INTO THE EFFECTS OF NONCOGNITIVE SKILLS ON HEALTH AND HUMAN CAPITAL ACCUMULATION

3.1 Introduction

The burgeoning literature on the economics of noncognitive skills has established the importance of these skills on wages, employment, smoking, crime, teen pregnancy, and tests of achievement (Heckman, Stixrud, and Urzua 2006; Borghans, Duckworth, Heckman, and ter Weel 2008). Noncognitive skills also have an important role in determining human capital accumulation in the form of schooling choices (Coleman and DeLeire 2003; Heckman, Stixrud, and Urzua 2006; Cebi 2007; Urzua 2008). The health literature has extensively documented the positive correlation between health and education (Grossman 2000; Grossman 2006). Both literatures discuss the need to incorporate noncognitive skills and health into one framework (Heckman 2007; Cuhna and Heckman 2007; Kaestner 2009). In an extended model of human capability formation Heckman (2008) suggests cognitive, noncognitive and health capabilities together contribute to various socioeconomic and adult outcomes. These literatures have just begun to address the effects of noncognitive skills on health and human capital accumulation, in the form of education (Kaestner 2009; Conti, Heckman, and Urzua 2010; Chiteji 2010). Both literatures remain silent about the effect of noncognitive skills on health and another form of human capital, human capital acquired on the job. One model of on-the-job human capital accumulation allows individuals to obtain human capital through a learning-by-doing technology that relies on hours of work (Shaw 1989; Imai and Keane 2004). Noncognitive skills such as locus of control and self-esteem affect hours of work (Urzua 2008). If noncognitive skills influence hours of work, then on-the-job human capital accumulation may depend on noncognitive skills.

This chapter puts the idea of health capital and on-the-job human capital from Chapter 1 together with the idea of noncognitive skills from Chapter 2 to explore the influence of these skills on human capital and health capital formation in adult life. A stock of noncognitive skills measured by the degree of future orientation, self-efficacy, trust-hostility, and aspirations is incorporated into a learning-by-doing model that allows for endogenous human capital and health capital accumulation. This stock of noncognitive skills exogenously enters the production of human capital and health capital. Individuals in this lifecycle model choose hours worked, medical consumption, nonmedical consumption subject to a budget constraint, a healthy time constraint, and the production functions for health and human capital. Exogenous health shocks directly affect sick time which then affect the choice of consumption and leisure. Estimation of the model's structural parameters follows a two-step procedure. The first step estimates the human capital and health capital production functions, allowing for noncognitive skills. The second step incorporates the production function parameters into the estimation of the utility parameters using a nonlinear generalized method of moments (GMM) estimator that comes directly from the solution to the individual's optimization problem.

The model uses data on working male head of households found in the Panel Study of Income Dynamics (PSID). The sample of men covering survey years 1989-2003 are between the ages of 37 and 60. The model relies on annual sick time and self-reported health status as measures of healthy time and the health stock. Measures of noncognitive skills come from questions asked 1968-1972. Medical out-of-pocket expenditures are imputed from the Consumer Expenditure Survey and enter the production of health.

The major findings show noncognitive skills do affect on-the-job human capital and but not health capital accumulation. Individuals with higher self-efficacy earn higher future wages. Since the wage represents the stock of human capital, individuals with higher self-efficacy produce higher levels of human capital. Individuals with higher trust-hostility earn lower future wages. None of the noncognitive skills exhibit any statistically significant relationship with future health. Utility parameter estimates support utility theory by suggesting diminishing returns to leisure and substitutability between consumption and leisure. In a specification with heterogeneity in education and demographics marginal utility of consumption is positive when evaluated at sample means. In addition, the mean value of -1.03 for the intertemporal substitution elasticity for consumption implies a 1 percent equiproportionate increase in all prices leads to a 1.03 percent reduction in consumption. Compared to a model without noncognitive skills in human capital or health production presented in Chapter 1, the results diminish the role of current health on future health.

3.2 Literature On Noncognitive Skills and Health

The noncognitive literature discusses the need to incorporate noncognitive skills and health into one framework (Heckman 2007; Cuhna and Heckman 2008). Heckman (2008) extends a conceptual model of human capability formation to include cognitive, noncognitive and health capabilities, suggesting these capabilities together contribute to various socioeconomic and adult outcomes. In this context these capabilities are complementary, that is, noncognitive capabilities can help the development of health and cognitive capabilities and vice versa. Noncognitive skills can help the formulation of human capital and health capital by affecting self regulation and choices (Cunha and Heckman 2007). One example is the rate of time preference. The rate of time preference can affect an individual's choice of healthy or unhealthy behaviors and investment in schooling.[28] Kaestner (2009) is one study about the effect of noncognitive and cognitive skills on health. Kaestner examines the association between cognitive skills and noncognitive skills measured at the end of childhood (age of 14 or 15) with adult health measured at the age of 41 in the National Longitudinal Survey of Youth 1979. Kaestner measures cognitive skills with the AFQT score and noncognitive skills with self-esteem, locus of control, church attendance, history of stealing, and drug use. Kaestner considers the effect of these skills on self-reported physical health and mental health by gender. For men and women adolescent cognitive ability and one noncognitive ability, self-esteem, are significantly associated with better health in adulthood. For men the magnitude of the direct associations is similar for cognitive ability and self-esteem. For women, though, the magnitude of the direct association is higher for self-esteem than cognitive ability.

Conti, Heckman, and Urzua (2010) study health disparities across education groups accounting for early cognitive, noncognitive, and health endowments. They estimate a structural model of schooling choice and its affect on post-schooling health, allowing for early cognitive, noncognitive, and health endowments to influence schooling choice and health. In this way, they can identify the true causal effect of education on health and estimate the entire distribution of this effect. They estimate the model with

[28] Fuchs (1982) proposed this time preference hypothesis. Fuchs argues individuals who are more future oriented, or have a higher degree of time preference, make larger investments in health and schooling.

data from the 1970 British Cohort Study, relating endowments at the age of 10 to labor market and health disparities at the age of 30. Noncognitive factors are measured with locus of control, perseverance, cooperativeness, completeness of tasks, attentiveness, and persistence. Cognitive factors are measured with tests on language comprehension, math, and reading. The health outcomes they consider include self-reported overall health, obesity, and depression, and the health behaviors they consider include smoking, exercise, and drug use. They find noncognitive factors substantially diminish the role of cognitive factors in predicting health outcomes, suggesting noncognitive factors are relatively more important in promoting health and healthy behaviors. Education plays a more important causal role for men than women in accounting for gaps in obesity rates and exercise. Furthermore, the effect of education on health varies by level of cognitive and noncognitve abilities. On most health outcomes for men, the benefit of education is much larger at the lower end of the noncognitive distribution and at the higher end of the cognitive distribution.

Chiteji (2010) also studies the relationship between noncognitive skills and healthy behavior, but instead uses the Panel Study of Income Dynamics (PSID) to ascertain the effect of self-efficacy and being future oriented on alcohol consumption and exercising. Self-efficacy is measured by a scale developed by the PSID and describes the evaluation of one's ability to effectively perform tasks necessary to achieve a goal. Being future oriented comes from a question asking the length of the individual's time horizon and characterizes an individual's rate of time preference. Both self-efficacy and time horizon are collected in 1972. Chiteji estimates logit models that relate noncognitive skills to alcohol consumption in 1972 and exercising in 1999. The results suggest individuals who are future oriented are less likely to drink and more likely to exercise. These effects are consistent with the idea that more patient individuals are less likely to engage in unhealthy behaviors and invest in health production. Self-efficacy has similar effects. Individuals with higher levels of self-efficacy face a lower likelihood of drinking and higher likelihood of exercising.

This chapter contributes to this literature by assessing the effect of noncognitive skills measured early in life on adult health. In addition, this chapter also assesses the effect of noncognitive skills on on-the-job human capital accumulation. The model presented in the next section allows noncognitive skills to simultaneously affect the production of human capital and health capital.

3.3 Model of Noncognitive Skills, Human Capital Accumulation, and Health Capital Accumulation

The model draws on elements from the health and labor literatures by combining health capital with human capital through a learning-by-doing technology. It allows health to enter utility directly and act as a constraint on behavior. Noncognitive skills enter the model in the same way as education enters the Grossman (1972) model. Like education noncognitive skills are taken as exogenous. In this model noncognitive skills enter the production of health and human capital. An individual maximizes a standard utility function that is additively separable over time and is defined over leisure (L_t), nonmedical consumption (C_t), and the stock of health (H_t). In each period,

$$U = U(L_t, C_t, H_t)$$

Income comes from two sources, asset income ($r_t A_t$) and labor income ($w_t N_t$). This income can be spent on nonmedical consumption and used to purchase medical services (M_t) at the price p_t^m. Putting income and expenditures together gives the following intertemporal budget constraint:

$$A_{t+1} = (1 + r_t)(A_t + w_t N_t - C_t - p_t^m M_t) \text{ (Asset Accumulation)}$$

The observed wage (w_t) is the product of a human capital stock (K_t) and the unobserved rental rate on human capital (R_t), so unlike the standard labor supply model the wage is endogenous.

$$w_t = R_t K_t \text{ (Wage Equation)}$$

In each period an individual inherits a stock of human capital which depreciates at rate δ_K. Human capital evolves in a similar way as health capital. Next period's human capital is the previous period's stock less depreciation plus new investment. New investment occurs on the job through learning-by-doing that depends on hours worked, the levels of human and health capital, and the exogenously determined stock of noncognitive skills, $x(N_t, K_t, H_t; Noncog)$. In this context noncognitive skills can raise the efficiency of human capital production. $Noncog$ is a stock of noncognitive skills characterized by personality traits that describe an individual's time horizon, self-efficacy, trust and hostility, and aspirations. Human capital evolves according to

$$K_{t+1} = (1 - \delta_K)K_t + x(N_t, K_t, H_t; Noncog)$$

$$= f(N_t, K_t, H_t; Noncog) \text{ (Human Capital Accumulation)}$$

In each period an individual inherits a stock of health capital (H_t) which depreciates at rate δ_H and can be replenished by devoting time to health, L_t (e.g. exercise), and purchasing medical services. Leisure and medical services enter a health production function, $y(M_t, L_t)$. The same stock of noncognitive skills that enter human capital production also enter health production exogenously and can raise the efficiency of health production, so $y(M_t, L_t; Noncog)$. Health capital then evolves according to

$$H_{t+1} = (1 - \delta_H)H_t + y(M_t, L_t; Noncog)$$
$$= I(M_t, L_t, H_t; Noncog) \text{(Health Capital Accumulation)}$$

Individuals also have healthy time (ht_t) which can be allocated across two activities: leisure (L_t) and work (N_t). The healthy time constraint becomes

$$L_t + N_t = ht_t \text{ (Healthy Time Constraint)}$$

In addition, total time (T) is the sum of healthy time (ht_t) and sick time (s_t), so

$$ht_t + s_t = T \text{ (Total Time Constraint)}$$

The individual's optimization problem can be represented in a dynamic programming framework with state variables,(A_t, K_t, H_t), and choice variables (C_t, L_t, M_t).

$$V(A_t, K_t, H_t) = \max_{C_t, L_t, M_t} \{U(L_t, C_t, H_t) + \beta V(A_{t+1}, K_{t+1}, H_{t+1})\}$$
$$s.t. A_{t+1} = (1 + r_t)(A_t + w_t N_t - C_t - p_t^m M_t) \text{ (Asset Accumulation)}$$

$$w_t = R_t K_t \text{ (Wage Equation)}$$

$$K_{t+1} = f(N_t, K_t, H_t; Noncog) \text{ (Human Capital Accumulation)}$$

$$H_{t+1} = I(M_t, L_t, H_t; Noncog) \text{(Health Capital Accumulation)}$$

$$L_t + N_t = ht_t \text{ (Healthy Time Constraint)}$$

$$ht_t + s_t = T \text{ (Total Time Constraint)}$$

Like other life cycle labor supply models the source of uncertainty in this model comes from unknown future realizations of tastes, prices, wages, and interest rates (MaCurdy 1983). This model does not make assumptions about the form of the distributions generating the uncertainty. Unlike other models of life cycle labor supply, this model adds uncertain health shocks through sick time. s_t also follows a distribution of unknown form. The individual in this model does not know when she will be sick. s_t characterizes acute illnesses that an individual cannot anticipate such as the common cold, food poisoning, and the flu.[29] s_t acts as an exogenous shock to the total time and healthy time constraints.[30]

The optimization problem can be simplified by substituting the wage equation and time constraints. With choice variables $C_{t,}$, N_t, and M_t the optimization problem can now be written as

$$V(A_t, K_t, H_t) = \max_{C_t, N_t, M_t} \{U(ht_t - N_t, C_t, H_t) + \beta V(A_{t+1}, K_{t+1}, H_{t+1})\}$$

$$s.t. A_{t+1} = (1 + r_t)(A_t + R_t K_t N_t - C_t - p_t^m M_t) \text{ (Asset Accumulation)}$$

$$K_{t+1} = f(N_t, K_t, H_t; Noncog) \text{ (Human Capital Accumulation)}$$

[29] Gilleskie (2010) also considers acute illnesses.
[30] This idea of s_t affecting the time constraint is similar in spirit to the literature on fixed time costs and labor supply (Cogan 1981).

$$H_{t+1} = I(M_t, ht_t - N_t, H_t; Noncog)(\text{Health Capital Accumulation})$$

The first order conditions with respect to C_t, N_t, and M_t:

$$C_t: U_{c,t} - \beta(1 + r_t)E_t\{V_A^{t+1}\} = 0 \quad (1)$$

$$N_t: -U_{L,t} + \beta E_t\{V_A^{t+1}(1 + r_t)R_t K_t + V_K^{t+1} f_{N,t} - V_H^{t+1} I_{L,t}\} = 0 \quad (2)$$

$$M_t: \beta E_t\{-V_A^{t+1}(1 + r_t)p_t^m + V_H^{t+1} I_{M,t}\} = 0 \quad (3)$$

A subscript represents a partial derivative with respect to that variable at time *t*. For example, $U_{c,t} = \frac{\partial U_t}{\partial C_t}$ and $U_{c,t+1} = \frac{\partial U_{t+1}}{\partial C_{t+1}}$. Equations 1 and 2 are most similar to conditions in life cycle labor supply models. Equation 1 is the standard Euler equation for consumption and describes the optimal consumption over time. Equation 2 shows the effect of endogenous human capital and health capital accumulation in the last two terms. Without endogenous human capital and health capital accumulation, these terms are zero, and Equation 2 reduces to a condition similar to the standard labor-leisure condition from the static labor supply model where $R_t K_t$ is replaced by w_t. Equation 3 describes the optimal amount of medical consumption. Noncognitive skills, suppressed to simplify notation, enter the model's first order conditions through the marginal utility and marginal product functions. The first order conditions can be combined with envelope conditions for the state variables (A_t, K_t, H_t) to solve the optimization problem. The following equilibrium condition characterizes the solution to the optimization problem.[31]

[31] See Appendix for derivation of solution to the optimization problem. With exogenous noncognitive skills the analytic solution to the model does not change. Only the specification of health and human capital production changes (discussed in section on Econometric Specification).

$$-U_{L,t} + \beta U_{c,t+1}(1+r_t)w_t + \beta f_{N,t} U_{c,t+1} R_{t+1} N_{t+1} - \frac{U_{c,t} p_t^m}{I_{M,t}} I_{L,t} +$$

$$\beta f_{N,t} \frac{f_{k,t+1}}{f_{N,t+1}} \left(U_{L,t+1} - U_{c,t+1} w_{t+1} + \frac{U_{c,t+1} p_{t+1}^m}{I_{M,t+1}} I_{L,t+1} \right) = 0 \quad (4)$$

Under rational expectations, eq(4) has zero expectation at time period *t*, so realizations of future variables imply

$$-U_{L,t} + \beta U_{c,t+1}(1+r_t)w_t + \beta f_{N,t} U_{c,t+1} R_{t+1} N_{t+1} - \frac{U_{c,t} p_t^m}{I_{M,t}} I_{L,t} +$$

$$\beta f_{N,t} \frac{f_{k,t+1}}{f_{N,t+1}} \left(U_{L,t+1} - U_{c,t+1} w_{t+1} + \frac{U_{c,t+1} p_{t+1}^m}{I_{M,t+1}} I_{L,t+1} \right) = u_{t+1} \quad (5)$$

where u_{t+1} is the forecast error at time *t*. Rational expectations implies $E_t\{u_{t+1}\} = 0$, so any information at time *t* is not useful in forecasting future variables. This orthogonality between u_{t+1} and the information at time *t* will be exploited to estimate the structural parameters of the model.

3.4 Econometric Specification

A two step estimation strategy is used to estimate the structural parameters of the model. In the first step quadratic human capital and health production function parameters are estimated with ordinary least squares, allowing for noncognitive skills. The second step incorporates these production function parameters in a nonlinear Generalized Method of Moments (GMM) estimation of the equilibrium condition (5) to identify utility parameters.

The specification for the human capital production function, $f(N_t, K_t, H_t; Noncog)$, is quadratic in its arguments. The quadratic specification represents the concave nature of earnings over the life-cycle:

$$K_{t+1} = f(N_t, K_t, H_t; Noncog) = \alpha_0 K_t + \alpha_1 K_t^2 + \alpha_2 K_t N_t + \alpha_3 N_t + \alpha_4 N_t^2 +$$
$$\alpha_5 H_t + \alpha_6 H_t^2 + \alpha_7 H_t N_t + \alpha_8 K_t H_t + \tau_t + \epsilon_{i,t}$$

The vector $\alpha = (\alpha_0, \alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6, \alpha_7, \alpha_8)$ contains the structural parameters of the human capital production function. Noncognitive skills enter the specification through varying slope parameters, so

$$\alpha_i = \alpha_{i0} + \alpha_{i1}\boldsymbol{D}\ i = 0,1,3, \dots, 6$$

Education and demographics enter similarly. ***D*** is a vector of education, demographic, and noncognitive skills, so ***D*={**education, nonwhite, horizon, self-efficacy, trust-hostility, and aspirations**}.** τ_t is an exogenous time-specific growth rate of human capital common to all individuals, and $\epsilon_{i,t}$ is an individual-specific component of human capital growth. As written, the human capital production function can not be directly estimated since the stock of human capital, K_t, is not observed. Wages are observed, so the specification can be rewritten in terms of wages using the relationship, $w_t = R_t K_t$. Making the substitution $K_t = \frac{w_t}{R_t}$ gives a specification that can be estimated

$$\frac{w_{t+1}}{R_{t+1}} = \alpha_0 \frac{w_t}{R_t} + \alpha_1 \left(\frac{w_t}{R_t}\right)^2 + \alpha_2 \left(\frac{w_t}{R_t}\right) N_t + \alpha_3 N_t + \alpha_4 N_t^2 +$$
$$\alpha_5 H_t + \alpha_6 H_t^2 + \alpha_7 H_t N_t + \alpha_8 \left(\frac{w_t}{R_t}\right) H_t + \tau_t + \epsilon_{i,t}$$

This specification of the wage equation relates future discounted wages to current hours worked, current discounted wages, current health, and the stock of noncognitive skills. It is estimated using ordinary least squares, setting $R_t = 1 \; \forall t$.

The marginal product of hours worked ($f_{N,t}$), marginal product of human capital stock ($f_{K,t}$), and the marginal product of health capital stock ($f_{H,t}$) come from differentiating the quadratic human capital production function

$$f_{N,t} = \alpha_2 \left(\frac{w_t}{R_t}\right) + \alpha_3 + 2\alpha_4 N_t + \alpha_7 H_t$$

$$f_{K,t} = \alpha_0 + 2\alpha_1 \frac{w_t}{R_t} + \alpha_2 N_t + \alpha_8 H_t$$

$$f_{H,t} = \alpha_5 + 2\alpha_6 H_t + \alpha_7 N_t + \alpha_8 \frac{w_t}{R_t}$$

Health capital production, $I(M_t, L_t, H_t; Noncog)$, follows a quadratic specification similar to human capital production. The health capital production is specified as

$$H_{t+1} = I(M_t, L_t, H_t; Noncog) = \theta_0 H_t + \theta_1 H_t^2 + \theta_2 H_t L_t + \theta_3 L_t + \theta_4 L_t^2 +$$

$$\theta_5 M_t + \theta_6 M_t^2 + \theta_7 H_t M_t + \theta_8 L_t M_t + \tau_t + v_{i,t}$$

where the vector $\theta = (\theta_0, \theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6, \theta_7, \theta_8)$ represents the structural parameters of health capital production. Noncognitive skills, education, and demographics enter the specification in a way similar to human capital production by varying slope parameters, so

$$\theta_j = \theta_{j0} + \theta_{j1}\boldsymbol{D} \; i = 0,1,3, \ldots ,6$$

τ_t is an exogenous time-specific growth rate of health capital common to all individuals, and $v_{i,t}$ is an individual-specific component of health capital growth. The corresponding marginal product of medical expenditures ($I_{M,t}$), leisure ($I_{L,t}$), and health stock ($I_{H,t}$) are

$$I_{M,t} = \theta_5 + 2\theta_6 M_t + \theta_7 H_t + \theta_8 L_t$$

$$I_{L,t} = \theta_2 H_t + \theta_3 + 2\theta_4 L_t + \theta_8 M_t$$

$$I_{H,t} = \theta_0 + 2\theta_1 H_t + \theta_2 L_t + \theta_7 M_t$$

The specification of the utility function follows a translog form. The translog utility function is quadratic in its arguments and represents a local second-order approximation to any utility function. It also does not impose the restrictions of additivity or homothecity associated with other common utility functions, such as CES or Cobb-Douglas (Christensen, Jorgenson, and Lau 1975). The exact specification is

$$U(L_t, C_t, H_t) = \gamma_0 \ln L_t + \gamma_1 \ln C_t + \gamma_2 (\ln L_t)(\ln C_t) + \gamma_3 (\ln L_t)^2 + \gamma_4 (\ln C_t)^2 +$$

$$\gamma_5 \ln H_t + \gamma_6 (\ln H_t)(\ln L_t) + \gamma_7 (\ln H_t)(\ln C_t) + \gamma_8 (\ln H_t)^2$$

The vector $\gamma = (\gamma_0, \gamma_1, \gamma_2, \gamma_3, \gamma_4, \gamma_5, \gamma_6, \gamma_7, \gamma_8)$ represents the structural parameters of the utility function. Identification of these parameters requires a normalization, so $\gamma_0 = 1.$ Due to the form of the equilibrium condition (5) γ_5 and γ_8 can not be identified so are set to zero. Differentiating the translog utility function gives the marginal utility of leisure ($U_{L,t}$) and the marginal utility of nonmedical consumption ($U_{C,t}$) as

$$U_{L,t} = \frac{1 + \gamma_2 \ln C_t + 2\gamma_3 \ln L_t + \gamma_6 \ln H_t}{L_t}$$

$$U_{C,t} = \frac{\gamma_1 + \gamma_2 \ln L_t + 2\gamma_4 \ln C_t + \gamma_7 \ln H_t}{C_t}$$

Substituting the corresponding marginal utility and marginal product functions into the equilibrium condition produces a highly nonlinear equation that can be parameterized by $\Gamma = (\alpha, \gamma, \theta, \beta)$. Γ is the vector of unknown population parameters. α and θ are estimated from the human capital and health capital production functions in the first step, and β is set to .95, leaving the utility parameters to be estimated in the second step. Let X_{it} be the vector of variables entering the *i*th individual's equilibrium condition in period *t*. The *i*th individual's equation can be represented by

$$f(X_{it}, \gamma; \alpha, \theta, \beta) = u_{it+1}$$

Rational expectations implies information in the information set Ω_{it} is not useful in forecasting future variables, so

$$E_t\{f(X_{it}, \gamma; \alpha, \theta, \beta) \cdot Z_{it}\} = 0$$

where Z_{it} contains elements of Ω_{it}. The orthogonality between $f(X_{it}, \gamma; \alpha, \theta, \beta)$ and Z_{it} is exploited to estimate γ in a GMM estimator. With panel data of *T* years for each individual population orthogonality conditions are derived by averaging over time,

$$E\frac{1}{T}\sum_{t=1}^{T}[f(X_{it}, \gamma; \alpha, \theta, \beta)Z_{it}] = E[M(X_i, Z_i, \gamma; \alpha, \theta, \beta)] = 0$$

The sample analogs of these population conditions are constructed by averaging over a random sample of *N* individuals, so

$$O_N(\gamma) = \frac{1}{N}\sum_{i=1}^{N}[M(X_i, Z_i, \gamma; \alpha, \theta, \beta)]$$

The GMM estimator of γ minimizes the quadratic form $O_N(\gamma)W_N O_N'(\gamma)$ or

$$\gamma = \underset{\gamma}{\operatorname{argmin}}\, O_N(\gamma)W_N O_N'(\gamma)$$

where the W_N is a symmetric positive definite weighting matrix. W_N is unknown so is replaced by $\widehat{W_N}$ constructed from the residuals of a nonlinear two stage least squares (NL2SLS) procedure, allowing for conditional heteroskedasticity.

3.5 Data Construction

The Panel Study of Income Dynamics (PSID) serves as the main data source for estimating the structural model with noncognitive skills. The PSID offers the advantage of covering a representative sample of U.S. individuals during the working portion of the life cycle. The PSID started with 4,800 families in 1968 and with efforts to follow all family members now contains over 7,000 families. Interviews occur annually for 1968-1996 and biennially for 1997-2007. The PSID consists of three samples: the original Survey Research Center (SRC) sample, the Survey of Economic Opportunity (SEO) sample, and the Latino sample. The original SRC sample represents the U.S. population while the SEO sample serves as a supplementary low income subsample. The Latino sample, added in 1990, accounts for the changing nature of immigration in the U.S. This chapter excludes the SEO and Latino samples. Individuals are drawn from the SRC sample from 1989-2003 who meet the following criteria: (1) in family at time of interview; (2) head of household; (2) male; (3) not self-employed; (4) employed; (5)work at least 3 years; (6) age between 37 and 60; (7) real annual food consumption between

$520 and one-third of real annual family income; (8) annual hours worked at least 100; and (9) real hourly wage between $2/hour and $200/hour. Limiting the sample to male head of households of working age avoids issues with labor force nonparticipation and joint labor supply decisions. The sample selection criteria produce an unbalanced panel. This analysis treats the panel as continuous for the period 1989-2003 (ignoring biennial interruptions) and considers missing person-year observations as missing conditionally at random.

The PSID collects a variety of information on the labor market and socioeconomic characteristics of each household. Questions about annual family income and annual hours worked refer to the previous calendar year.[32] Food consumption also refers to the previous calendar year. Total annual food consumption includes food at home, food delivered to the home, eating out, and the value of food stamps. Total food consumption, deflated by the food component of the Consumer Price Index (2008 base year), is the consumption measure in the model. The PSID collects food consumption each survey year with the exception of survey years 1988 and 1989. For survey years after 1993, the PSID stopped reporting annual values for food at home, food delivered, eating out, and the value of food stamps. For these years annual amounts were created by annualizing the amount reported according to the reporting period (daily, weekly, biweekly, monthly, and monthly).

The hourly wage rate depends on whether the worker is hourly or salaried. For hourly workers the PSID gives the reported wage rate. For salaried workers the PSID

---

[32] Annual family income with negative values after 1993 is set to 1 to match the PSID bottom coding of family income in previous years.

reports an hourly rate that depends on the salary and the pay period (weekly, biweekly, or monthly). For these workers the PSID constructs a wage adjusted for a fixed number of hours each pay period instead of using the actual hours worked. For example, a salaried worker paid weekly has an hourly rate that is the salary divided by 40 hours. Salaried workers paid biweekly, monthly, and yearly have salaries divided by 80, 160, and 2,000 hours, respectively. All wage and income data are deflated by the personal consumption deflator using a 2008 base year.

Questions about annual sick hours and self-rated health serve as measures of healthy time and the health stock, respectively. PSID respondents report the amount of work missed due to own illness for the previous year. Specifically, the PSID asks "Did you miss work because you were sick? How much work did you miss?" The PSID calculates the annual sick hours as weeks ill times 8 hours for the first 8 weeks and times 60 hours for any weeks thereafter. Beginning in 1994 the PSID did not calculate annual sick hours, so annual sick hours are created by applying the same formula to the number of reported days, weeks, or months missed due to own sickness. The PSID self-rated health question began in 1984 and asks "Would you say your health in general is excellent, very good, good, fair, or poor?"[33] These responses are converted to a four point scale in the following way: 1 (Fair or Poor); 2 Good; 3 Very Good; and 4 Excellent.

The PSID collected a series of attitudinal questions from 1969-1972 that were later dropped from the survey. The PSID reports indexes that cover four personality traits and represent the stock of noncognitive skills in the model: (1) horizon length, (2)

[33] The self-rated health question in the PSID has been used in the recent health literature. Fletcher and Sindelar (2009) and Fletcher, Sindelar, and Yamaguchi (2009) are examples.

self-efficacy, (3) trust-hostility, and (4) aspirations. Each trait is measured by an index comprised of questions asked to head of households. The index value for each trait in the analysis is the average over the years 1969-1972. The index for horizon length ranges from 0 to 8 and describes the degree of future orientation. It is based on the following statements:

1. Is sure whether will or will not move
2. Has explicit plans for children's education
3. Has plans for an explicit kind of new job
4. Knows and mentions what kind of training new job requires
5. Has substantial savings relative to income
6. Has definite expectations that next few years will be better or worse
7. Expects to have a child more than one year hence, or expects no more children and is doing something to limit the number of children

The index for self-efficacy ranges from 0 to 7 and describes one's ability to effectively perform tasks necessary to achieve a goal. It is based on the following statements:

1. Sure life would work out
2. Plans life ahead
3. Gets to carry out things
4. Finishes things
5. Rather save for future
6. Has no limitations
7. Thinks about things that might happen in future

The index for trust-hostility ranges from 0 to 5 and is based on the following statements:

1. Does not get angry easily
2. Matters what others think
3. Trusts most other people
4. Believes life of average man getting better
5. Believes not a lot of people have good things they don't deserve

The index for aspirations and ambition ranges from 0-8 and describes an individual's degree of aspirations and ambition. It is based on the following statements:

1. Might make purposive move
2. Wanted more work and/or worked >2500 hours
3. Might quit a job because it was not challenging
4. Prefers a job with chances for making more money even if dislikes job
5. Spends time figuring out how to get more money
6. Plans to get a new job, and knows what type of job, and knows what it might pay
7. Plans for job regardless of details

Other variables necessary for the model are the interest rate and medical out-of-pocket expenditures. The interest rate in the model reflects the after-tax annual 3-month Treasury bill interest rate based on a marginal federal tax rate from the NBER TAXSIM module.[34] Medical out-of-pocket expenditures are imputed using data from the Consumer Expenditure Survey. The appendix explains the imputation procedure for medical out-of-pocket expenditures.

---

[34] The marginal federal tax rate comes from the NBER TAXSIM module using head of household, labor income, and number of children as input variables.

After applying the sample selection criteria, the analysis sample contains 2,525 person-year observations and covers 1989-2003. The questions for noncognitive skills asked in 1968-1972 limit the sample to head of households who are typically not from split-off families. As a consequence, the sample will be considerably older and unhealthier than the analysis sample from Chapter 1. Since instrument sets using *t-2* values are used in estimation, the sample size will also be considerably smaller. Table 3.1 displays summary statistics. The sample consists of males, who are mostly white and have completed almost 14 years of education. These men earn an average hourly wage of $27.75/hour and work about 2,245 hours each year. The average age is 49 years old, over 10 years older than the average age in the sample from Chapter 1. With an average health status at 2.82, health in the sample is almost "Very Good," which is less than the average health reported in the sample from Chapter 1. Individuals take almost 33 hours of sick time each year and spend over $3,300 on medical out-of-pocket expenditures.

Tables 3.2 and Table 3.3 report average horizon, self-efficacy, trust-hostility, and aspiration scores by the main outcome variables, wage and health status. In addition, Table 3.4 reports the average scores by education level. Men in the lowest quartile of the wage distribution tend to have the shortest time horizon and the lowest levels of self-efficacy and trust-hostility. Men in the top quartile of the wage distribution report the longest time horizon and highest levels of self-efficacy and trust-hostility. However, men in the bottom quarter of the wage distribution have the highest level of aspirations. Levels of aspiration decrease with the wage. When moving from Poor or Fair health to Excellent health, health exhibits a positive relationship with each noncognitive skill, so men with excellent health have the longest time horizon and highest self-efficacy, trust-hostility, and aspirations. The same positive relationship holds with education for horizon, self-efficacy, and trust-hostility. Men who have completed college or more exhibit the longest time horizon and highest self-efficacy and trust-hostility. Education and aspiration do not follow this relationship. The average aspiration score increases from less than high school to high school and college then falls with men who have completed college or more.

Several variables must be scaled to facilitate the computation of the model. Leisure, hours worked, and healthy time are divided by 1,000. Total food consumption and medical out-of-pocket expenditures are divided by 10,000. The hourly wage and age are divided by 100.

3.6 Results

Table 3.5 and Table 3.6 show the estimation of the human capital and health capital production functions using ordinary least squares. Heteroskedasticity robust standard errors are in parentheses. Columns 1 and 2 of each table provide parameter estimates for the base specifications without heterogeneity while columns 3 and 4 introduce education, demographic, and noncognitive heterogeneity. This interaction with education, nonwhite, each noncognitive skill allows shifts in the production functions. Only the specification without heterogeneity shows a concave relationship between current wages and future wages. Since the wage represents the human capital stock, the results show human capital production increases with the current stock of human capital at a decreasing rate, suggesting decreasing marginal productivity of human capital. The specification without heterogeneity suggests for a given level of human capital, health improves human capital accumulation. The specification with heterogeneity suggests for a given level of human capital, schooling augments the production of post-schooling human capital, or the two are complements. The positive interaction with wage and hours of work suggests hours of work also improves human capital production, or hours of work and human capital are complements. Among the noncognitive parameters of interest, only self-efficacy and trust-hostility exhibit a statistically significant relationship with the future wage through the current wage. For a given level of the current wage, individuals with higher self-efficacy earn more, or produce higher levels of human capital. The interaction with the quadratic wage suggests a concave relationship between self-efficacy and wage. For a given level of the current wage, individuals with higher trust-hostility earn less at an increasing rate. Aspirations have a negative effect through the quadratic wage. Trust-hostility has a positive effect through health.

The health capital specification without heterogeneity suggests future health capital stock is an increasing function of current health. This relationship between current health and future health disappears after introducing heterogeneity. In fact, none of the noncognitive skills exhibit any statistically significant relationship with future health. For a given level of current leisure, being nonwhite lowers future health.

Comparing the production parameter estimates from this model with noncognitive skills to the production parameter estimates from the model without noncognitive skills presented in Chapter 1 (Table 1.2 and Table 1.3) shows some differences. The human capital production function with heterogeneity presented Table 1.2 exhibits diminishing returns to human capital while the human capital production function with noncognitive skills does not exhibit diminishing returns to human capital. After adding noncognitive skills, education becomes statistically significant. The health capital production function with heterogeneity in Table 1.3 shows the importance of current health for future health. The importance of current health for future health diminishes after adding noncognitive skills.

Table 3.7 displays the nonlinear generalized method of moments estimation of utility parameters. Columns 1 and 2 are based on the estimated production parameters without heterogeneity while columns 3 and 4 are based on production parameters that allow heterogeneity. The coefficient on leisure is set to 1 for identification of the remaining parameters. The instrument set used in the estimation includes time *t*, time *t-1*, and time *t-2* values of leisure, food consumption, wage, health status, medical expenditures, the after-tax interest rate, age, education, region dummies, nonwhite, number of children, family size, horizon, self-efficacy, trust-hostility, aspirations, annual time dummies, interactions between leisure and food consumption, wage and health status, health status and leisure, medical expenditures and health status, leisure and medical expenditures, and age and education. Squared values at time *t*, time *t-1*, and time *t-2* of food consumption, wage, health status, medical expenditures, age, and leisure are also included.

The specification without heterogeneity shows positive parameter estimates for consumption and leisure. The negative sign on the interaction between consumption and leisure indicates they are substitutes in preferences. When evaluated at the mean values of the data, marginal utility of consumption and leisure are negative which is possible with translog utility.[35] Leisure exhibits strong diminishing returns. The interaction

---

[35] When evaluated at sample means, marginal utility of consumption is -0.001 and marginal utility of leisure is -0.006.

between consumption and health suggest they are complements. The test statistics for the Sargan test suggests the model specification without heterogeneity can be rejected at the 10 percent level of significance, but the model specification with heterogeneity can not be rejected. Adding education and demographic heterogeneity preserves all signs except the quadratic term for consumption, the interaction between health and leisure, and the interaction between health and consumption. Marginal utility of consumption is positive, and marginal utility of leisure is negative when evaluated at the means of the data.[36] Education lowers utility through leisure and consumption. The mean value of -1.03 (without heterogeneity) for the intertemporal substitution elasticity for consumption implies a 1 percent equiproportionate increase in all prices leads to a 1.03 percent reduction in consumption.

Comparing the utility parameter estimates from this model with noncognitive skills in the production functions to the analogous model without noncognitive skills presented in Chapter 1 (Table 1.6) shows similarities and differences. The specifications without heterogeneity in utility parameters are similar in signs with the exception of quadratic consumption and the interactions between health and leisure and health and consumption. The specifications with demographic and education heterogeneity only differ in signs on the heterogeneity parameters for leisure and the and health and consumption interaction.

3.7 Conclusion and Future Work

This chapter synthesizes the idea of health and on-the-job human capital accumulation from Chapter 1 with the idea of noncognitive skills in Chapter 2 to examine

[36] When evaluated at sample means, marginal utility of consumption is 0.001 and marginal utility of leisure is -0.003.

the influence of these skills on human capital accumulation and health capital accumulation in adult life. While the interaction of noncognitive skills with health and education is just receiving attention, the interaction of noncognitive skills with health and on-the-job human capital has not been explored in the noncognitive, health, or human capital literatures. This chapter presents an exploratory analysis by introducing noncognitive skills to a life cycle labor supply model with endogenous health and human capital accumulation. Noncognitive skills, measured by degree of future orientation, self-efficacy, trust-hostility, and aspirations, are allowed to exogenously affect human capital production and health production in the same way as education exogenously affects health production in the Grossman (1972) model of health. The model's structural parameters are estimated using nonlinear generalized method of moments. The model takes advantage of noncognitive skills assessed in the early years of the Panel Study of Income Dynamics and relates these skills to health and human capital accumulation during adult life.

The main findings suggest a role for noncognitive skills on on-the-job human capital accumulation. Self-efficacy positively affects future wages, or human capital accumulation. Individuals with high self-efficacy receive higher future wages. Individuals with higher trust-hostility receive lower future wages. Unlike the literature relating noncognitive skills to better health, this analysis does not find any evidence that noncognitive skills influence future health. The finding that noncognitive skills aid the development of on-the-job human capital is new and underscores the importance of developing these skills during the early part of life.

Future work should explore other noncognitive measures in the PSID. The current model incorporates a few of many attitudinal and personality questions available in the PSID. Other information can be included in the model such as the K-6 Psychological Distress Scale which measures emotional stability. Individual components of the horizon, self-efficacy, trust-hostility, and aspirations indexes can be included in the model in place of each index for robustness. Chapter 2 reports differences in wages by race and gender due to noncognitive skills. An analysis by race and gender should be pursued. The recent noncognitive literature suggests complementarities between cognitive and noncognitive skills. Interactions between education, a measure of cognitive skills, and noncognitive skills can be included to test for any nonseparabilities between the two sets of skills.

Table 3.1 Summary Statistics 1989-2003

Variable (n=2,525)	Mean	Std. Dev.	Min.	Max.
Wage (2008 dollars)	27.75	16.81	2.38	145.84
Work hours	2244.48	517.00	150	5436
Sick hours	32.77	119.29	0	1980
Leisure hours	6482.76	512.84	3276	8610
Healthy hours	8727.23	119.29	6780	8760
Health status	2.82	0.89	1	4
Food Consumption (2008 dollars)	9591.87	4331.72	743.51	38989.45
Medical Out-of-pocket (2008 dollars)	3323.70	984.97	53.58	12959.17
After-tax interest rate	3.42	0.76	1.70	7.34
Age	49.42	4.82	37	60
Education (years)	13.95	2.41	6	17
Number of children	0.59	0.90	0	5
Family Size	2.91	1.19	1	8
Northeast	0.17	0.37	0	1
North Central	0.37	0.48	0	1
South	0.29	0.45	0	1
West	0.17	0.37	0	1
White	0.93	0.25	0	1
Nonwhite	0.07	0.25	0	1
Married	0.88	0.33	0	1
Horizon	4.98	0.88	2	7.25
Self-Efficacy	4.06	1.25	0.50	7
Trust-Hostility	2.60	1.07	0	5
Aspirations	3.30	1.29	0.25	7.5

Table 3.2 Average Noncognitive Skills by Wage Percentile

Wage	Horizon	Self-Efficacy	Trust-Hostility	Aspirations
Less than 25th percentile	4.79	3.63	2.38	3.34
Between 25th and 75th percentiles	4.94	4.10	2.62	3.29
Greater than 75th percentile	5.23	4.37	2.76	3.27

Table 3.3 Average Noncognitive Skills by Health Status

Health Status	Horizon	Self-Efficacy	Trust-Hostility	Aspirations
Poor or Fair	4.67	3.54	2.26	3.17
Good	4.97	3.95	2.60	3.35
Very Good	5.00	4.07	2.63	3.24
Excellent	5.06	4.35	2.66	3.37

Table 3.4 Average Noncognitive Skills by Education

Education	Horizon	Self-Efficacy	Trust-Hostility	Aspirations
Less than high school	4.71	3.60	2.12	3.17
High school and college	4.87	3.86	2.52	3.34
Greater than college	5.19	4.44	2.82	3.28

Table 3.5 Human Capital Production Parameters

Parameter	(1) Estimate	(2) Std. Error	(3) Estimate	(4) Std. Error
α_0 (w_t)	0.883***	(0.220)	0.349	(0.297)
α_0 (w_t) (educ)			0.0546***	(0.0178)
α_0 (w_t) (nonwhite)			-0.186**	(0.0738)
α_0 (w_t) (horizon)			-0.0328	(0.0547)
α_0 (w_t) (self-efficacy)			0.0757**	(0.0330)
α_0 (w_t) (trust-hostility)			-0.163***	(0.0411)
α_0 (w_t) (aspirations)			0.0342	(0.0283)
α_1 (w_t^2)	-0.680***	(0.102)	-0.306	(0.352)
α_1 (w_t^2) (educ)			-0.0517**	(0.0252)
α_1 (w_t^2)(nonwhite)			0.147*	(0.0819)
α_1 (w_t^2)(horizon)			0.0718	(0.0743)
α_1 (w_t^2) (self-efficacy)			-0.117***	(0.0417)
α_1 (w_t^2) (trust-hostility)			0.267***	(0.0487)
α_1 (w_t^2) (aspirations)			-0.0657**	(0.0331)
α_2 (w_tN_t)	0.144	(0.0895)	0.182***	(0.0403)
α_3 (N_t)	0.00518	(0.0247)	0.0894	(0.0673)
α_3 (N_t)(educ)			0.00356	(0.00353)
α_3 (N_t)(nonwhite)			-0.0284	(0.0182)
α_3 (N_t)(horizon)			-0.0143	(0.00965)
α_3 (N_t)(self-efficacy)			-0.00956	(0.0109)
α_3 (N_t) (trust-hostility)			-0.0125	(0.0116)
α_3 (N_t) (aspirations)			-0.00115	(0.00803)
α_4 (N_t^2)	-0.00295	(0.00361)	-0.0229	(0.0148)
α_4 (N_t^2)(educ)			-0.00148*	(0.000861)
α_4 (N_t^2)(nonwhite)			0.00473	(0.00380)
α_4 (N_t^2)(horizon)			0.00476**	(0.00213)
α_4 (N_t^2) (self-efficacy)			0.00231	(0.00212)
α_4 (N_t^2) (trust-hostility)			0.00205	(0.00247)
α_4 (N_t^2) (aspirations)			0.000744	(0.00179)
α_5 (H_t)	-0.00572	(0.0174)	-0.0260	(0.0635)
α_5 (H_t)(educ)			-0.00570	(0.00348)
α_5 (H_t)(nonwhite)			0.0554***	(0.0212)
α_5 (H_t)(horizon)			0.00569	(0.00920)
α_5 (H_t)(self-efficacy)			0.00194	(0.00941)
α_5 (H_t)(trust-hostility)			0.0228**	(0.00955)
α_5 (H_t)(aspirations)			0.000518	(0.00795)
α_6 (H_t^2)	0.000207	(0.00276)	0.00549	(0.0121)
α_6 (H_t^2)(educ)			0.000980	(0.000727)
α_6 (H_t^2)(nonwhite)			-0.0107***	(0.00402)
α_6 (H_t^2)(horizon)			-0.000545	(0.00179)
α_6 (H_t^2)(self-efficacy)			-0.000510	(0.00168)
α_6 (H_t^2)(trust-hostility)			-0.00341**	(0.00174)
α_6 (H_t^2)(aspirations)			-0.000819	(0.00152)
α_7 (H_tN_t)	-0.00450	(0.00568)	-0.00339	(0.00575)
α_8 (w_tH_t)	0.0732**	(0.0349)	0.0115	(0.0297)
Observations	2525		2525	
R-squared	0.920		0.932	

Regressions include annual time dummy variables. Robust standard errors in parentheses.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$ Columns 3 and 4 include observed heterogeneity.

Table 3.6 Health Capital Production Parameters

Parameter	(1) Estimate	(2) Std. Error	(3) Estimate	(4) Std. Error
θ_0 (H_t)	0.498**	(0.234)	0.443	(0.699)
θ_0 (H_t) (educ)			-0.0149	(0.0356)
θ_0 (H_t) (nonwhite)			0.334	(0.394)
θ_0 (H_t) (horizon)			0.0148	(0.0914)
θ_0 (H_t) (self-efficacy)			0.0490	(0.0754)
θ_0 (H_t) (trust-hostility)			0.0375	(0.0765)
θ_0 (H_t) (aspirations)			-0.0366	(0.0774)
θ_1 (H_t^2)	-0.0169	(0.0168)	0.0314	(0.121)
θ_1 (H_t^2)(educ)			0.00463	(0.00680)
θ_1 (H_t^2)(nonwhite)			-0.0991	(0.0764)
θ_1 (H_t^2)(horizon)			-0.00909	(0.0172)
θ_1 (H_t^2)(self-efficacy)			-0.00749	(0.0139)
θ_1 (H_t^2)(trust-hostility)			-0.0115	(0.0142)
θ_1 (H_t^2)(aspirations)			-0.000158	(0.0141)
θ_2 (H_tL_t)	0.0241	(0.0312)	0.0347	(0.0338)
θ_3 (L_t)	-0.0982	(0.284)	-0.864	(0.557)
θ_3 (L_t)(educ)			0.0164	(0.0246)
θ_3 (L_t)(nonwhite)			-0.401*	(0.231)
θ_3 (L_t)(horizon)			0.0668	(0.0620)
θ_3 (L_t)(self-efficacy)			-0.00377	(0.0500)
θ_3 (L_t)(trust-hostility)			-0.0112	(0.0519)
θ_3 (L_t)(aspirations)			0.0363	(0.0462)
θ_4 (L_t^2)	-0.00349	(0.0212)	0.0635	(0.0562)
θ_4 (L_t^2)(educ)			-0.00223	(0.00287)
θ_4 (L_t^2)(nonwhite)			0.0487**	(0.0247)
θ_4 (L_t^2)(horizon)			-0.00707	(0.00687)
θ_4 (L_t^2)(self-efficacy)			4.56e-05	(0.00501)
θ_4 (L_t^2)(trust-hostility)			0.00160	(0.00554)
θ_4 (L_t^2)(aspirations)			-0.00220	(0.00507)
θ_5 (M_t)	0.519	(1.776)	1.210	(6.179)
θ_5 (M_t)(educ)			0.0954	(0.274)
θ_5 (M_t)(nonwhite)			2.758	(3.242)
θ_5 (M_t)(horizon)			-0.574	(0.749)
θ_5 (M_t)(self-efficacy)			-0.117	(0.602)
θ_5 (M_t)(trust-hostility)			-0.0336	(0.538)
θ_5 (M_t)(aspirations)			-0.482	(0.509)
θ_6 (M_t^2)	-1.392**	(0.580)	-7.895	(8.482)
θ_6 (M_t^2)(educ)			0.00558	(0.425)
θ_6 (M_t^2)(nonwhite)			-4.855	(5.580)
θ_6 (M_t^2)(horizon)			0.784	(1.043)
θ_6 (M_t^2)(self-efficacy)			0.0876	(0.839)
θ_6 (M_t^2)(trust-hostility)			-0.0552	(0.685)
θ_6 (M_t^2)(trust-hostility)			1.064	(0.732)
θ_7 (H_tM_t)	0.259	(0.189)	-0.0681	(0.233)
θ_8 (L_tM_t)	0.0984	(0.253)	0.466	(0.352)
Observations	2525		2525	
R-squared	0.948		0.950	

Regressions include annual time dummy variables. Robust standard errors in parentheses.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$ Columns 3 and 4 include observed heterogeneity.

Table 3.7 Utility Parameters Using Time t, t-1, and t-2 Instruments

Parameter	(1) Estimate	(2) Std. Error	(3) Estimate	(4) Std. Error
γ_0 (ln L_t)	1.00		1.00	
γ_0 (ln L_t) (educ)			-0.0110***	(0.0007)
γ_0 (ln L_t) (nonwhite)			0.0314	(0.2171)
γ_1 (ln C_t)	0.0474***	(0.0045)	0.0115***	(0.0011)
γ_1 (ln C_t) (educ)			-0.0002***	(0.00001)
γ_1 (ln C_t) (nonwhite)			0.0003	(0.0024)
γ_2 (ln L_t) (ln C_t)	-0.0274***	(0.0029)	-0.0052***	(0.0006)
γ_3 (ln L_t)^2	-0.2796***	(0.0123)	-0.2296***	(0.0034)
γ_4 (ln C_t)^2	0.0003	(0.0004)	-0.0062	(0.0121)
γ_6 (ln H_t) (ln L_t)	0.0039	(0.0335)	-0.0143***	(0.0045)
γ_7 (ln H_t) (ln C_t)	0.0030***	(0.0013)	-0.0002***	(0.00004)
ISE (mean)	-1.03		0.14	
ISE (median)	-0.77		-0.05	
Sargan test [df]	120.14		112.15[96]	
(p-value)	[100] (0.08)		(0.12)	
Observations	2525		2525	

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1
Columns 3 and 4 include observed heterogeneity.

Appendix: Derivation of Equilibrium Condition and Estimating Equation

The individual's optimization problem is

$$
\begin{aligned}
V(A_t, K_t, H_t) &= \max_{C_t, L_t, M_t} \{U(L_t, C_t, H_t) + \beta V(A_{t+1}, K_{t+1}, H_{t+1})\} \\
s.t.\, A_{t+1} &= (1 + r_t)(A_t + w_t N_t - C_t - p_t^m M_t) \text{ (Asset Accumulation)} \\
w_t &= R_t K_t \text{ (Wage Equation)} \\
K_{t+1} &= (1 - \delta_K) K_t + x(N_t, K_t, H_t) \\
&= f(N_t, K_t, H_t) \text{ (Human Capital Accumulation)} \\
H_{t+1} &= (1 - \delta_H) H_t + y(M_t, L_t) = I(M_t, L_t, H_t) \text{(Health Capital Accumulation)} \\
L_t + N_t &= ht_t \text{ (Healthy Time Constraint)} \\
ht_t + s_t &= T \text{ (Total Time Constraint)}
\end{aligned}
$$

The Total Time Constraint is extraneous. The Total Time Constraint can be used to define $ht_t = T - s_t$, and the Wage Equation can be substituted into the Asset Accumulation Equation. The optimization problem can now be written as

$$
\begin{aligned}
V(A_t, K_t, H_t) &= \max_{C_t, N_t, M_t} \{U(ht_t - N_t, C_t, H_t) + \beta V(A_{t+1}, K_{t+1}, H_{t+1})\} \\
s.t.\, A_{t+1} &= (1 + r_t)(A_t + R_t K_t N_t - C_t - p_t^m M_t) \text{ (Asset Accumulation)} \\
K_{t+1} &= f(N_t, K_t, H_t) \text{ (Human Capital Accumulation)} \\
H_{t+1} &= I(M_t, ht_t - N_t, H_t) \text{(Health Capital Accumulation)} \\
L_t + N_t &= ht_t \text{ (Healthy Time Constraint)}
\end{aligned}
$$

Substitute for L_t from the Healthy Time Constraint ($L_t = ht_t - N_t$) gives

$$
\begin{aligned}
V(A_t, K_t, H_t) &= \max_{C_t, N_t, M_t} \{U(ht_t - N_t, C_t, H_t) + \beta V(A_{t+1}, K_{t+1}, H_{t+1})\} \\
s.t.\, A_{t+1} &= (1 + r_t)(A_t + R_t K_t N_t - C_t - p_t^m M_t) \text{ (Asset Accumulation)} \\
K_{t+1} &= f(N_t, K_t, H_t) \text{ (Human Capital Accumulation)} \\
H_{t+1} &= I(M_t, ht_t - N_t, H_t) \text{(Health Capital Accumulation)}
\end{aligned}
$$

First order conditions with respect to C_t, N_t, and M_t:

$$C_t\colon U_{c,t} - \beta(1 + r_t) E_t\{V_A^{t+1}\} = 0 \;\Rightarrow\; U_{c,t} = \beta(1 + r_t) E_t\{V_A^{t+1}\} \text{ (6)}$$

$$N_t\colon -U_{L,t} + \beta E_t\left\{V_A^{t+1}(1 + r_t) R_t K_t + V_K^{t+1} f_{N,t} - V_H^{t+1} I_{L,t}\right\} = 0 \;\Rightarrow \text{(7)}$$

$$-U_{L,t} + \beta E_t\{V_A^{t+1}(1 + r_t) R_t K_t\} + \beta E_t\left\{V_K^{t+1} f_{N,t}\right\} - \beta E_t\left\{V_H^{t+1} I_{L,t}\right\} = 0$$

$$M_t\colon \beta E_t\left\{-V_A^{t+1}(1+r_t)p_t^m + V_H^{t+1}I_{M,t}\right\} = 0 \quad (8)$$

$$-\beta E_t\{V_A^{t+1}(1+r_t)p_t^m\} + \beta E_t\left\{V_H^{t+1}I_{M,t}\right\} = 0$$

Envelope conditions:

$$A_t\colon V_A^t = \beta E_t\{(1+r_t)V_A^{t+1}\} \quad (9)$$

$$K_t\colon V_K^t = \beta E_t\left\{V_A^{t+1}(1+r_t)R_tN_t + V_K^{t+1}f_{k,t}\right\} \ (10)$$

Substitute eq(6) into eq(8) for $E_tV_A^{t+1} = \frac{U_{c,t}}{\beta(1+r_t)}$ gives

$$E_t\left\{-U_{c,t}p_t^m\right\} + \beta E_t\left\{V_H^{t+1}I_{M,t}\right\} = 0 \Rightarrow E_t\, V_H^{t+1} = \frac{U_{c,t}p_t^m}{\beta I_{M,t}} \quad (11)$$

Substitute eq(8) into eq(7) for $E_t\, V_H^{t+1}$:

$$-U_{L,t} + \beta E_t\{V_A^{t+1}(1+r_t)R_tK_t\} + \beta\left\{V_K^{t+1}f_{N,t}\right\} - E_t\left\{\frac{U_{c,t}p_t^m}{\beta I_{M,t}}I_{L,t}\right\} = 0$$

Substitute for V_A^{t+1} and V_K^{t+1} after updating envelope conditions (eq(9) and eq(10))

$$-U_{L,t} + \beta E_t\{\beta E_{t+1}\{(1+r_{t+1})V_A^{t+2}\}(1+r_t)R_tK_t\} +$$

$$\beta E_{t+1}\left\{\beta E_t\left\{V_A^{t+2}(1+r_{t+1})R_{t+1}N_{t+1} + V_K^{t+2}f_{k,t+1}\right\}f_{N,t}\right\} - \frac{U_{c,t}p_t^m}{\beta I_{M,t}}I_{L,t} = 0 \quad (12)$$

Update eq(6) to substitute for $\beta E_{t+1}\{(1+r_{t+1})V_A^{t+2} = U_{c,t+1}$:

$$-U_{L,t} + \beta E_t\left\{U_{c,t+1}(1+r_t)R_tK_t\right\} + \beta E_t\left\{U_{c,t+1}R_{t+1}N_{t+1} + \beta E_{t+1}V_K^{t+2}f_{k,t+1}\right\}f_{N,t}$$

$$-\frac{U_{c,t}p_t^m}{\beta I_{M,t}}I_{L,t} = 0 \ (13)$$

Update eq(7) to solve for $\beta E_{t+1} V_K^{t+2}$:

$$-U_{L,t+1} + \beta E_{t+1}\{V_A^{t+2}(1+r_{t+1})R_{t+1}K_{t+1}\} + \beta E_{t+1}\{V_K^{t+2} f_{N,t+1}\} - \beta E_{t+1}\{V_H^{t+2} I_{L,t+1}\} = 0 \tag{14}$$

Before solving for $\beta E_{t+1} V_K^{t+2}$, substitute for $\beta E_{t+1}\{(1+r_{t+1})V_A^{t+2}\} = U_{c,t+1}$ and $E_{t+1}\,\{V_H^{t+2}\} = \frac{U_{c,t+1} p_{t+1}^m}{\beta I_{M,t+1}}$ in eq(14)

$$-U_{L,t+1} + U_{c,t+1} R_{t+1} K_{t+1} + \beta E_{t+1}\{V_K^{t+2} f_{N,t+1}\} - \frac{U_{c,t+1} p_{t+1}^m}{\beta I_{M,t+1}} I_{L,t+1} = 0$$

Solve for $\beta E_{t+1}\{V_K^{t+2}\}$

$$\beta E_{t+1}\{V_K^{t+2}\} = \left[U_{L,t+1} - U_{c,t+1} R_{t+1} K_{t+1} + \frac{U_{c,t+1} p_{t+1}^m}{\beta I_{M,t+1}} I_{L,t+1}\right] \frac{1}{f_{N,t+1}}$$

Substitute for $\beta E_{t+1}\{V_K^{t+2}\}$ into eq(13)

$$-U_{L,t} + \beta E_t\{U_{c,t+1}(1+r_t) R_t K_t\} +$$

$$\beta E_t \left\{ U_{c,t+1} R_{t+1} N_{t+1} + \left(U_{L,t+1} - U_{c,t+1} R_{t+1} K_{t+1} + \frac{U_{c,t+1} p_{t+1}^m}{\beta I_{M,t+1}} I_{L,t+1} \right) \frac{f_{k,t+1}}{f_{N,t+1}} \right\} f_{N,t} - \frac{U_{c,t} p_t^m}{\beta I_{M,t}} I_{L,t} = 0$$

Substitute $w_t = R_t K_t$ and $w_{t+1} = R_{t+1} K_{t+1}$

$$-U_{L,t} + \beta E_t\{U_{c,t+1}(1+r_t) w_t\} +$$

$$\beta f_{N,t} E_t \left\{ U_{c,t+1} R_{t+1} N_{t+1} + \left(U_{L,t+1} - U_{c,t+1} w_{t+1} + \frac{U_{c,t+1} p_{t+1}^m}{\beta I_{M,t+1}} I_{L,t+1} \right) \frac{f_{k,t+1}}{f_{N,t+1}} \right\} - \frac{U_{c,t} p_t^m}{\beta I_{M,t}} I_{L,t} = 0$$

Zero expectation at time *t* implies the following estimating equation.

$$-U_{L,t} + \beta U_{c,t+1}(1+r_t) w_t + \beta f_{N,t} U_{c,t+1} R_{t+1} N_{t+1} - \frac{U_{c,t} p_t^m}{\beta I_{M,t}} I_{L,t} +$$

$$\beta f_{N,t} \frac{f_{k,t+1}}{f_{N,t+1}} \left(U_{L,t+1} - U_{c,t+1} w_{t+1} + \frac{U_{c,t+1} p_{t+1}^{m}}{\beta I_{M,t+1}} I_{L,t+1} \right) = 0 \ (15)$$

Eq(15) is the estimating equation.

Appendix: Consumption Imputation

Consumption is imputed based on a procedure developed by Blundell, Pistaferri, and Preston (2006) that imputes nondurable expenditures in the PSID using data from the CE. Blundell, Pistaferri, and Preston (2006, 2008) estimate a demand equation for food at home in the CE that depends on demographics, food prices, and nondurable expenditures. Their model specification follows:

$$\ln\left(c_{i,t}^{f}\right) = X_{i,t}\varphi + \pi \ln\left(c_{i,t}\right) + \epsilon_{i,t}$$

where $c_{i,t}^{f}$ is food at home expenditures and $c_{i,t}$ is nondurable expenditures. In their estimation Blundell, Pistaferri, and Preston instrument for nondurable expenditures to account for potential measurement error. With estimates of $\hat{\varphi}$ and $\hat{\pi}$ from the CE along with demographics and food expenditures from the PSID, nondurable consumption can be predicted in the PSID using

$$\widehat{\ln(c_{i,t})} = \frac{\ln\left(c_{i,t}^{f}\right) - X_{i,t}\hat{\varphi}}{\hat{\pi}}$$

As long as food at home expenditures are monotonic in nondurable expenditures, the estimates from the food demand equation, $\hat{\varphi}$ and $\hat{\pi}$, are estimated consistently, and the trends in the variance of actual and imputed are similar, the imputation procedure will give consistent estimates of nondurable expenditures in the PSID. For this chapter imputed consumption comes from a simpler version of Blundell, Pistaferri, and Preston (2006). Nondurable expenditures are predicted using

$$\widehat{\ln(c_{i,t})} = \frac{\ln\left(c_{i,t}^{f}\right) - 3.6674 - .5746 \ln\left(cpi_{i}^{f}\right)}{.4573}$$

Luigi Pistaferri kindly provided this estimation.

Appendix: Medical Out-of-Pocket Expenditures Imputation

Medical out-of-pocket expenditures (MOOP) are imputed using the Consumer Expenditure Survey (CE). Conducted by the Census Bureau for the Bureau of Labor Statistics the CE collects extensive information on the buying habits of American households since 1980. The Census Bureau currently uses the CE to impute out-of-pocket medical expenditures in the Current Population Survey for its experimental poverty measure (Betson 2001; O'Donnell and Beard 2009).[37] The CE consists of two components: the weekly Diary Survey and the quarterly Interview Survey. The Diary Survey follows households for two consecutive weeks and records expenditures on small, frequently purchased items such as food, personal care, and household supplies. The Interview Survey follows households for five consecutive quarters and collects large expenditures on a quarterly basis, including property, automobiles, and appliances. The Interview Survey also collects various medical expenditures, so it is used for the out-of-pocket medical expenditure imputation. Specifically, the Interview Survey records expenditures on six medical categories: (1) drug preparations; (2) ophthalmic products and orthopedic appliances; (3) physicians, dentists, and other medical professionals; (4) hospitals; (5) nursing homes; and (6) health insurance. The Interview Survey sample is selected on a rotating panel basis each quarter. Approximately 7,000 households that are a representative sample of the US population are followed for five quarters (U.S. Bureau of Labor Statistics 2008).

[37] The Census Bureau's two part model separately imputes MOOP for individuals with zero medical expenditures and individuals with positive medical expenditures. The model draws on data from the 1996-1997 CE and “ages” the imputed data using the Consumer Price Index. The variables in the imputation include age, income as a percent of poverty, health insurance coverage, and family size (O'Donnell and Beard 2009).

Ed Harris and John Sabelhaus of the Congressional Budget Office maintain CE extract files through the National Bureau of Economic Research.[38] Harris and Sabelhaus provide "family-level" extract files that aggregate four quarterly records into an annual record for each family and collapse the hundreds of spending, income, and wealth categories into a consistent set of categories across all years (Harris and Sabelhaus 2000). Their out-of-pocket medical expenditure category sums the six medical spending categories. Because of the rotating panel design of the Interview Survey, each family's annual record covers the spending for the first quarter the family entered the sample and the following three quarters. For example, a family entering the survey the first quarter of 1997 reports expenditures for the 1997 calendar year, January 1997 through December 1997. A family entering the survey the second quarter of 1997 reports expenditures for April 1997 through March 1998. To match the calendar year timing of the PSID, only the quarter one CE extract files are used.

Harris and Sabelhaus (2000) recommend selecting a usable CE sample that meet the following criteria. First, the respondent must meet the BLS "complete income reporter" requirement. Second, the household must have completed four quarterly interviews. Third, the household should not be a student household. Additionally, to best match the PSID sample, the CE sample is limited to male head of households between the ages of 25 and 60. These requirements produce a CE sample of 8,321 member observations covering 1980-2003.

The imputation procedure involves estimating the following annual regression on CE variables that are also available in the PSID.

---

[38] Available at http://www.nber.org/data/ces_cbo.html.

$$medical_i = \beta_0 + \beta_1 age_i + \beta_2 famsize_i + \beta_3 yrs\_educ_i + \beta_4 income_i + \beta_5 northeast_i$$
$$+\beta_6 midwest_i + \beta_7 south_i + \beta_8 white_i + \epsilon_i$$

$medical_i$ is the sum of six medical spending categories. age_i is the age of the individual. $famsize_i$ is family size. yrs_educ_i is the individual's years of education corrected to match the PSID definition of education. $income_i$ is a broad family income measure that includes (1) wages and salaries; (2) self-employment income; (3) rent, interest, and dividends; (4) government transfers; and (5) rent received as pay. Negative values of $income_i$ are set to 1 to match bottom coding of the PSID family income variable. $northeast_i$, $midwest_i$, and $south_i$ are region dummy variables while $white_i$ is a dummy variable.

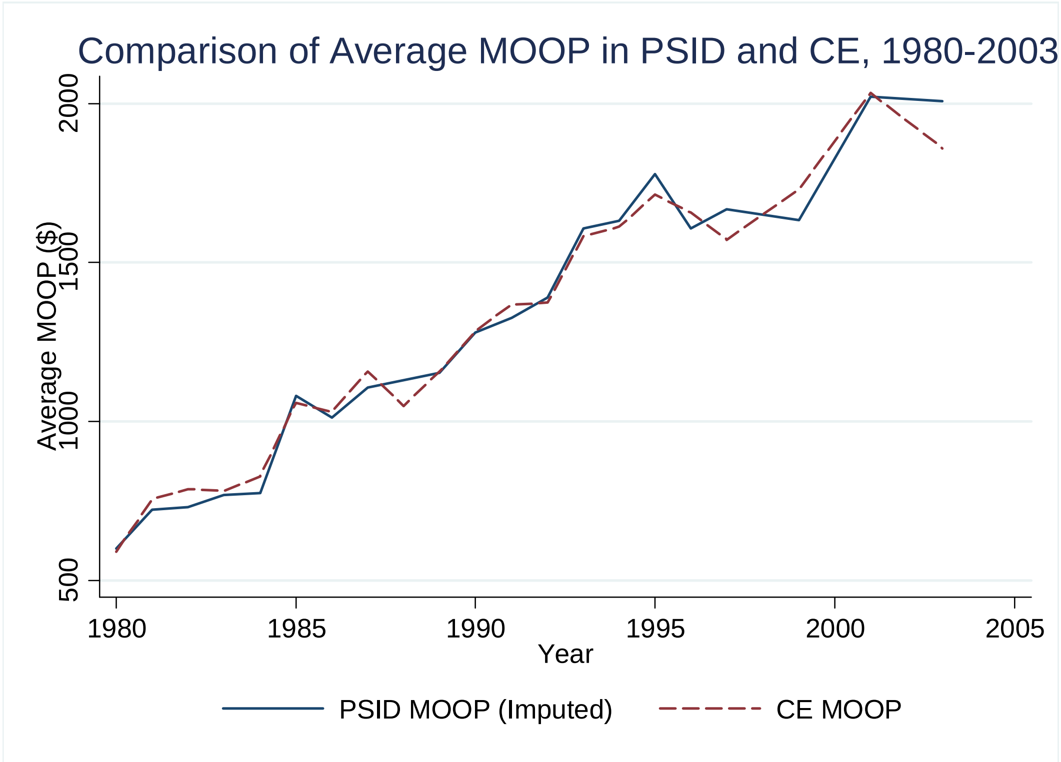
The imputation regressions omit the top one percent of medical expenditures to mitigate the effect of outliers. The imputation regressions also omit 1988 because the education variable is zero for all observations in 1988. After estimating the regression on CE data, the annual coefficients are applied to PSID data to predict out-of-pocket medical expenditures. The predicted out-of-pocket medical expenditures are deflated using the medical component of the consumer price index with a 2008 base year.

Table A.1 and the accompanying figures compare the actual medical expenditures in the CE to the imputed medical expenditures in the PSID before deflating. Table A.1 shows the average of the imputed variable exceeds the average of the actual variable while the variance of the imputed variable is smaller. Comparing averages over time shows the imputed variable average seems to match the actual variable average over the sample period (Figure A.1, top panel). The standard deviation of the actual variable always exceeds the standard deviation of the imputed variable (Figure A.1, bottom panel). Both standard deviation series exhibit similar trends over time. The kernel density figures show both variables have similar distributions (Figure A.2).

Table A.1 CE and PSID Comparison of Medical Out-of-Pocket Expenditures (MOOP)

Variable	Mean	Std. Dev.
CE MOOP	1350.41	1414.17
PSID MOOP (Imputed)	1459.35	704.62

Figure A.0.1 MOOP Trends



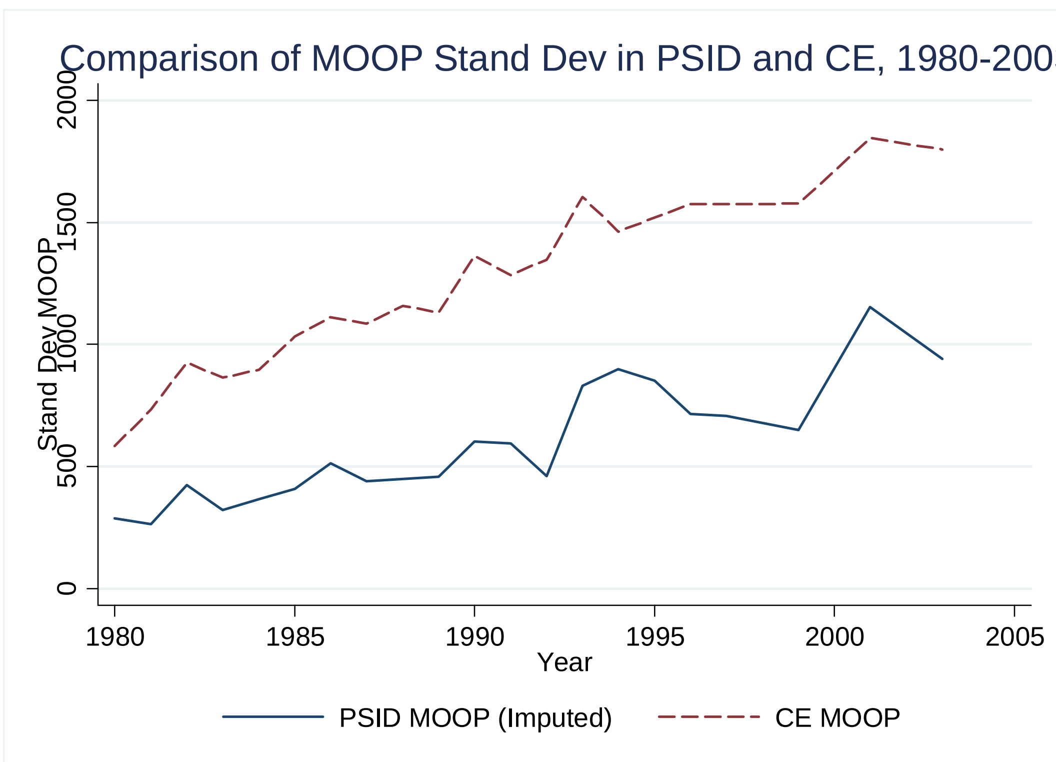
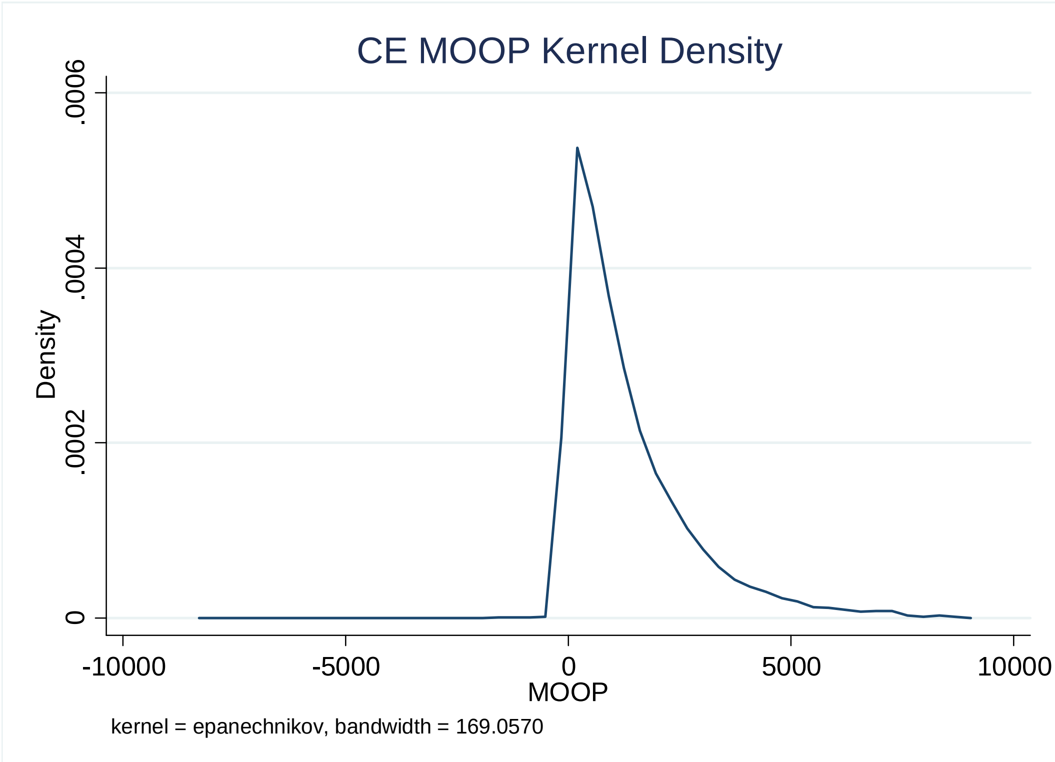
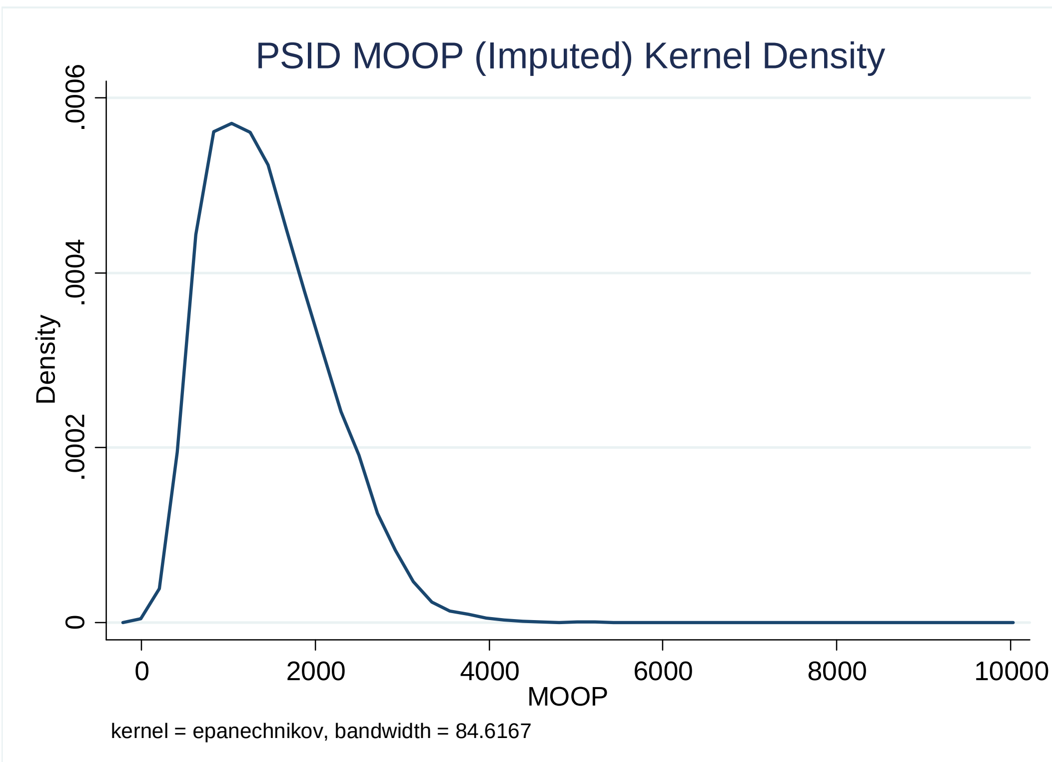


Figure A.0.2 MOOP Kernel Density





References
Chapter One References

Altonji, Joseph. 1986. Intertemporal Substitution in Labor Supply: Evidence from Micro Data. *Journal of Political Economy* 94, Part 2, S176-S215.

Becker, Gary. 1965. A Theory of the Allocation of Time. *Economic Journal* 75, 493-517.

Ben-Porath, Yoram. 1967. The Production of Human Capital and the Life Cycle of Earnings. *Journal of Political Economy* 75, 352-365.

Berger, Mark, and Belton Fleisher. 1984. Husband's Health and Wife's Labor Supply. *Journal of Health Economics* 3, 63-75.

Berger, Mark. 1983. Labor Supply and Spouse's Health: The Effects of Illness, Disability and Mortality. *Social Science Quarterly* 64, 494-509.

Betson, David. 2001. Imputation of Medical Out of Pocket (MOOP) Spending to CPS Records. Census Bureau Technical Report.

Blundell, Richard, Luigi Pistaferri, and Ian Preston. 2008. Consumption Inequality and Partial Insurance. *American Economic Review* 98, no. 5, 1887-1921.

Blundell, Richard, Luigi Pistaferri, and Ian Preston. 2006. Imputing Consumption in the PSID Using Food Demand Estimates from the CEX. Working Paper, University College London Department of Economics.

Bound, John, Todd Stinebrickner, and Timothy Waidmann. 2010. Health, Economic Resources and the Work Decisions of Older Men. *Journal of Econometrics* 156, 106-129.

Bound, John, Michael Schoenbaum, Todd Stinebrickner, and Timothy Waidmann. 1999. The Dynamic Effects of Health on the Labor Force Transitions of Older Workers. *Labour Economics* 6, 179-202.

Card, David. 1999. The Causal Effects of Education on Earnings. Chapter 30, Handbook of Labor Economics, Vol. 3A, 1801-1863.

Christensen, Laurits, Jorgenseon, Dale, and Lawrence Lau. 1975. Transcendental Logarithmic Utility Functions. *American Economic Review* 65, no. 3, 367-383.

Cogan, John. 1981. Fixed Costs and Labor Supply. *Econometrica* 49, no. 4, 945-963.

Currie, Janet, and Brigitte Madrian. 1999. Health, Health Insurance, and the Labor Market. Chapter 50, Handbook of Labor Economics, Vol. 3, 3309-3416.

Erbsland, Manfred, Walter Ried, and Volker Ulrich. 1995. Health, Health Care, and the Environment. Econometric Evidence from German Micro Data. *Health Economics* 4, 169-182.

Fletcher, Jason, and Jody Sindelar. 2009. Estimating Causal Effects of Early Occupational Choice on Later Health: Evidence Using the PSID. NBER Working Paper No. 15256.

Fletcher, Jason, Sindelar, Jody, and Shintaro Yamaguchi. 2009. Cumulative Effects of Job Characteristics on Health. NBER Working Paper No. 15121.

French, Eric. 2005. The Effects of Health, Wealth, and Wages on Labour Supply and Retirement Behavior. *Review of Economic Studies* 72, 395-427.

Gilleskie, Donna. 2010. Work Absences and Doctor Visits During An Illness Episode: The Differential Role of Preferences, Production, and Policies Among Men and Women. *Journal of Econometrics* 156, 148-163.

Grossman, Michael. 2006. Education and Nonmarket Outcomes. Chapter 10, Handbook of the Economics of Education, Vol. 1, 578-633.

Grossman, Michael. 2000. The Human Capital Model. Chapter 7, Handbook of Health Economics, Vol. 1, 348-408.

Grossman, Michael. 1972. On the Concept of Health Capital and the Demand for Health. *Journal of Political Economy* 80, 223-255.

Hansen, Lars, and Kenneth Singleton. 1982. Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models. *Econometrica* 50, 1269-86.

Harris, Ed, and John Sebelhaus. 2000. Consumer Expenditure Survey. Congressional Budget Office.

Heckman, James. 1976. A Life Cycle Model of Earnings, Learning, and Consumption. *Journal of Political Economy* 84, S11-44.

Heckman, James, Lance Lochner, and Petra Todd. 2006. Earnings Functions, Rates of Return, and Treatment Effects: The Mincer Equation and Beyond. Chapter 7, Handbook of Education Economics, Vol. 1, 307-458.

Hotz, Joseph, Finn Kydland, and Guilheme Sedlacek. 1988. Intertemporal Preferences and Labor Supply. *Econometrica* 56, no. 2, 335-360.

Hum, Derek, Wayne Simpson, and Eyob Fissuh. 2008. "The Impact of Health on Labour Supply Over the Lifecycle." paper presented to joint sessions of the Canadian Labour

Market and Skills Researcher Network (CLSRN) and the Canadian Economics Association, Vancouver, June 5, 2008, and to the Oxford Business and Economics Conference 2008, Oxford, June 23, 2008.

Hum, Derek, Wayne Simpson, and Eyob Fissuh Ghebretsadik. 2006. The Impact of Health On Labour Supply in Panel Data. *Global Business and Economics Anthology*, Vol.1, 185-196.

Imai, Susumu, and Michael Keane. 2004. Intertemporal Labor Supply and Human Capital Accumulation. *International Economic Review* 45, no. 2, 601-641.

MaCurdy, Thomas. 1983. A Simple Scheme For Estimating An Intertemoral Model of Labor Supply and Consumption in the Presence of Taxes and Uncertainty. *International Economic Review* 24, no. 2, 265-289.

MaCurdy, Thomas. 1981. An Empirical Model of Labor Supply in a Life-Cycle Setting. *Journal of Political Economy* 88, 1059-85.

O'Donnell, Sharon, and Rodney Beard. 2009. Imputing Medical Out of Pocket (MOOP) Expenditures Using SIPP and MEPS. Census Bureau Technical Report.

Polachek, Solomon. 2008 Earnings Over the Life Cycle: The Mincer Earnings Function and Its Applications. Foundations and Trends in Microeconomics, 4(3), 165-272.

Ruhm, Christopher. 2000. Are Recessions Good For Your Health? *Quarterly Journal of Economics* 115, no. 2, 617-650.

Shaw, Kathryn. 1989. Life-Cycle Labor Supply with Human Capital Accumulation. *International Economic Review* 30, no. 2, 431-56.

Sickles, Robin, and Abdo Yazbeck. 1998. On the Dynamics for Leisure and the Production of Health. *Journal of Business and Economic Statistics* 16, no. 2, 187-197.

U.S. Bureau of Labor Statistics. 2008. Consumer Expenditure Survey Anthology, 2008. U.S. Department of Labor Report 1009.

Wagstaff, Adam. 1986. The Demand For Health: Some New Empirical Evidence. *Journal of Health Economics* 5, 195-233.

Weiss, Yoram. 1972. On the Optimal Pattern of Labor Supply. Economic Journal 82, 1292-1315.

Chapter Two References

Andrisani, Paul. 1977. Internal-External Attitudes, Personal Initiative, and the Labor Market Experience of White and Black Men. *Journal of Human Resources* 12, no. 3:308-28.

Bollinger, Christopher. 2003. Measurement Error in Human Capital and the Black-White Wage Gap. *Review of Economics and Statistics* 85, no.3, 578-585.

Bowles, Samuel, and Herbert Gintis. 1976. Schooling in capitalist America:Educational reform and the contradictions of economic life. New York: Basic Books.

Carneiro, Pedro, James J. Heckman, and Dimitriy V. Masterov. 2005. Labor market discrimination and racial differences in pre-market factors. *Journal of Law and Economics* 48, no.1, 1–39.

Cebi, Merve. 2007. Locus of Control and Human Capital Investment Revisited. *Journal of Human Resources* 42, no. 4, 919-932.

Chandra, Amitabh. 2003. Is the Convergence in the Racial Wage Gap Illusory? NBER Working Paper 9476.

Coleman, Margo, and Thomas DeLeire. 2003. An Economic Model of Locus of Control and the Human Capital Investment Decision. *Journal of Human Resources* 38, no. 3, 701-721.

Duncan, Greg, and James Morgan. 1981. Sense of Efficacy and Changes in Economic Status—A Comment on Andrisani. *Journal of Human Resources* 16, no. 4, 649-57.

Edwards, Richard C. 1976. Individual traits and organizational incentives:What makes a "good" worker? *Journal of Human Resources* 11, no. 1, 51–68.

Goldsmith, Arthur, William A.Darity, and Jonathan Veum. 1997. The impact of psychological and human capital on wages. *Economic Inquiry* 35, no. 4, 818–29.

Heckman, James. 2007. The Economics, Technology, and Neuroscience of Human Capability Formation. *Proceedings of the National Academy of Sciences* 104, no. 33, 13250-13255.

Heckman, James, Jora Stixrud, and Sergio Urzua. 2006. The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior. *Journal of LaborEconomics* 24, no. 3, 411-482.

Krueger, Alan, and David Schkade. 2008. Sorting in the Labor Market: Do Gregarious Workers Flock to Interactive Jobs? *Journal of Human Resources* 43, no. 4, 859–883.

Maxwell, Nan. 1994. The Effect on Black-White Wage Differences in the Quantity and Quality of Education. *Industrial and Labor Relations Review* 47, no. 2, 249-264.

Mueser, Peter. 1979. The Effects of Noncognitive Traits. Chapter 5 in Who Gets Ahead? The Determinants of Economic Success in America. by Christopher Jencks et. al. New York: Basic Books.

National Longitudinal Survey of Youth. 2010. Wages. Online documentation. Section 4.40, Accessed December 2010.

National Longitudinal Survey of Youth. 2007. Attitudes and Expectations. Online documentation Section 4.5, Accessed December 2007.

Neal, Derek. 2004. The Measured Black-White Wage Gap among Women is Too Small. *Journal of Political Economy* 112, no. 1, pt. 2, S1-S28.

Neal, Derek, and William Johnson. 1996. The Role of Premarket Factors in Black-White Wage Differences. *Journal of Political Economy* 104, no. 5, 869-895.

O'Neill, June. 1990. The Role of Human Capital in Earnings Differences Between Black and White Men. Journal of Economic Perspectives 4, no. 4, 25-45.

Rouse, Katy. 2008. The Impact of High School Leadership on Subsequent Educational Attainment. University of North Carolina-Chapel Hill Department of Economics Working Paper.

ter Weel, Bas. 2008. The Noncognitive Determinants of Labor Market and BehavioralOutcomes. *Journal of Human Resources* 43, no. 4, 729–737.

Urzua, Sergio. 2008. Racial Labor Market Gaps: The Role of Abilities and Schooling Choices. *Journal of Human Resources*. 43, no. 4, 919–971.

Weinberger, Catherine. 2008a. Who Do You Want on Your Team? Sports Skills, Math Skills, and Outcomes over the Life Cycle. American Economic Association Meetings, January 2008.

Weinberger, Catherine. 2008b. Quantifying the Growing Importance of "Soft Skills" in the Labor Market. American Economic Association Meetings, January 2008.

Chapter Three References

Borghans, Lex, Angela Lee Duckworth, James Heckman, and Bas ter Weel. 2008. The Economics and Psychology of Personality Traits. *Journal of Human Resources*. 43, no. 4, 972-1059.

Cebi, Merve. 2007. Locus of Control and Human Capital Investment Revisited. *Journal of Human Resources* 42, no. 4, 919-932.

Chiteji, Ngina. 2010. Time Preference, Noncognitive Skills and Well Being Across the Life Course: Do Noncognitive Skills Encourage Healthy Behavior? *American Economic Review Papers and Proceedings* 100, no. 2, 200-204.

Christensen, Laurits, Jorgenseon, Dale, and Lawrence Lau. 1975. Transcendental Logarithmic Utility Functions. *American Economic Review* 65, no. 3, 367-383.

Cogan, John. 1981. Fixed Costs and Labor Supply. *Econometrica* 49, no. 4, 945-963.

Coleman, Margo, and Thomas DeLeire. 2003. An Economic Model of Locus of Control and the Human Capital Investment Decision. *Journal of Human Resources* 38, no. 3, 701-721.

Conti, Gabriela, James Heckman, and Sergio Urzua. 2010. The Education-Health Gradient. *American Economic Review Papers and Proceedings* 100, no. 2, 234-238.

Cuhna, Flavio, and James Heckman. 2008. Formulating, Identifying and Estimating the Technology Cognitive and Noncognitive Skill Formation. *Journal of Human Resources* 43, no. 4, 738-782.

Cuhna, Flavio, and James Heckman. 2007. The Economics of Human Development: TheTechnology of Skill Formation. *American Economic Review Papers and Proceedings* 97, no. 2, 31-47.

Fletcher, Jason, and Jody Sindelar. 2009. Estimating Causal Effects of Early Occupational Choice on Later Health: Evidence Using the PSID. NBER Working Paper No. 15256.

Fletcher, Jason, Sindelar, Jody, and Shintaro Yamaguchi. 2009. Cumulative Effects of Job Characteristics on Health. NBER Working Paper No. 15121.

Fuchs, V.R. 1982. “Time Preference and Health: An Exploratory Study”, in: V.R. Fuchs, ed., Economic Aspects of Health (University of Chicago Press for the National Bureau of Economic Research, Chicago) 93-120.

Gilleskie, Donna. 2010. Work Absences and Doctor Visits During An Illness Episode: The Differential Role of Preferences, Production, and Policies Among Men and Women. *Journal of Econometrics* 156, 148-163.

Grossman, Michael. 2006. Education and Nonmarket Outcomes. Chapter 10, Handbook of the Economics of Education, Vol. 1, 578-633.

Grossman, Michael. 2000. The Human Capital Model. Chapter 7, Handbook of Health Economics, Vol. 1, 348-408.

Grossman, Michael. 1972. On the Concept of Health Capital and the Demand for Health. *Journal of Political Economy* 80, 223-255.

Heckman, James. 2008. Schools, Skills, and Synapses. *Economic Inquiry* 46, no. 3, 289-324.

Heckman, James. 2007. The Economics, Technology, and Neuroscience of Human Capability Formation. *Proceedings of the National Academy of Sciences* 104, no. 33, 13250-13255.

Heckman, James, Jora Stixrud, and Sergio Urzua. 2006. The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior *Journal of LaborEconomics* 24, no. 3, 411-482.

Imai, Susumu, and Michael Keane. 2004. Intertemporal Labor Supply and Human Capital Accumulation. *International Economic Review* 45, no. 2, 601-641.

Kaestner, Robert. 2009. Adolescent Cognitive and Non-cognitive Correlates of Adult Health. NBER Work Paper 14924, April 2009.

MaCurdy, Thomas. 1983. A Simple Scheme For Estimating An Intertemoral Model of Labor Supply and Consumption in the Presence of Taxes and Uncertainty. *International Economic Review* 24, no. 2, 265-289.

Shaw, Kathryn. 1989. Life-Cycle Labor Supply with Human Capital Accumulation. *International Economic Review* 30, no. 2, 431-56.

Urzua, Sergio. 2008. Racial Labor Market Gaps: The Role of Abilities and Schooling Choices. *Journal of Human Resources*. 43, no. 4, 919–971.

Vita

Charles Hokayem

Birthdate/Place: January 6, 1978, Gainesville, Georgia

Education

Master of Science, Economics, University of Wisconsin-Madison, August 2005

Bachelor of Science, Cum Laude, Economics and Mathematics, Centre College, May 2000

Professional Research Experience

Graduate Research Fellow, August 2007-Present, University of Kentucky Center for Poverty Research

Research Assistant, June 2006-July 2007, Center for Business and Economic Research, University of Kentucky

Senior Economic Research Associate, June 2001-August 2003, Federal Reserve Bank of St. Louis

Associate Economist, June 2000-May 2001, RTI International

Academic Awards and Honors

Graduate Research Fellow, University of Kentucky Center for Poverty Research, 2007-Present

Research Challenge Trust Fund Gatton Doctoral Fellowship, University of Kentucky, 2006-2007

National Institute of Mental Health Fellowship/Traineeship, University of Wisconsin, 2003-2005

Phi Beta Kappa, Centre College, 2000

The Harry Landreth Prize in Economics, John C. Young Scholar, The Wall Street Journal Award, Centre College, 2000

Omicron Delta Epsilon, Centre College, 1998

Publications

"The Economic Impact of Child Care Subsidies For Kentucky." (with James Ziliak and Bradley Hardy) University of Kentucky Center for Poverty Research Occasional Policy Brief #3, December 2008.

"Child Care Subsidies and the Economic Well-Being of Recipient Families: A Survey and Implications for Kentucky." (with James Ziliak and Bradley Hardy) Technical Report to Kentucky Youth Advocates and the Kentucky Cabinet for Health and Family Services, June 2008.

"Childhood Influences on Youth Obesity." (with Tim Classen) *Economics and Human Biology*, 3(2), pp. 165-187, July 2005. (Lead article in special issue on obesity)

"Prospecting Among the Inquiries: An Enrollment Forecasting Model." (with Robert Martin, John Perry, and Jed Leaf) in Managerial Analysis and Decision Support, NACUBO Accounting Committee, eds., Washington, DC: National Association of College and University Business Officers, 2004.

"Deflation, Corrosive and Otherwise." (with James Bullard) Federal Reserve Bank of St. Louis *Economic Synopses*, No. 17, 2003.

"Deflation, Corrosive and Otherwise." (with James Bullard) Federal Reserve Bank of St. Louis *National Economic Trends*, July 2003.

"Economists Weigh All the Costs of Health Care." Federal Reserve Bank of St. Louis *Regional Economist*, July 2002.

"Why the Fed Should Ignore the Stock Market." (with James Bullard) Federal Reserve Bank of St. Louis *Central Banker*, Spring 2002.

"Prospecting Among the Inquiries." (with Robert Martin, John Perry, and Jed Leaf) *Business Officer*, January 2002.

"Assessment of Provider, Laboratory, and Family Costs of Screening Young Children for Elevated Blood Lead Levels." (with Amanda Honeycutt and Heather Carter-Young) Technical Report to Centers for Disease Control and Prevention, RTI International, October 2000.

Working Papers

"Health, Human Capital Accumulation, and Life Cycle Labor Supply" (with James Ziliak)

"Noncognitive Skills and the Racial Wage Gap"

"An Exploration into the Effects of Noncognitive Skills on Health and Human Capital Accumulation"

"Is Wetter Better?: Local Alcohol Policy and Economic Development" (with Ken Troske)

Professional Affiliations

American Economic Association, Southern Economic Association

Teaching Experience

Instructor, Summer 2007-Summer 2008, University of Kentucky

Teaching Assistant, 2005-2006, University of Kentucky