ABSTRACT

Title of dissertation: ESSAYS ON THE ECONOMICS OF ABILITY, EDUCATION AND LABOR MARKET OUTCOMES

María Fernanda Prada
Doctor of Philosophy, 2014

Dissertation directed by: Professor Sergio Urzúa
Department of Economics

The analysis of the heterogeneity in worker ability and its economic implications have been a focus of a broad strand of research in labor economics.

Several studies have demonstrated that both cognitive and socio-emotional dimensions of ability have a positive effect on wages, schooling, and the probability of choosing high paying occupations. However, there is no theoretical reason to expect that all dimensions affect outcomes in the same direction.

This dissertation, composed by four chapters, shows that mechanical ability, jointly with cognitive and socio-emotional dimensions, affects schooling decisions and labor market outcomes. Moreover, it demonstrates that this facet of ability has a positive economic return and affects schooling decisions and occupational choices differently than other measures of ability.

Chapter 2 introduces the concept of mechanical ability, describes the tests used to measure it, and briefly compares this dimension with conventional measures of ability.

Chapter 3 presents a general framework to understand the effects of multiple dimensions of ability on outcomes with special emphasis in the selection into occupations and tasks where workers are more productive. This framework is used to decompose the overall effect of unobserved abilities into the components explained by schooling decision, occupational choice, and direct on-the-job productivity. I show that all three dimensions of ability have multiple, heterogeneous, and independent roles. They influence the sorting of workers into schooling and occupations, and also have a direct effect on wages. This implies that a policy that increases ability at advanced ages, when schooling and occupational decisions cannot be altered, may still have a direct impact on wages.

Chapter 4, written in collaboration with Sergio Urzúa, analyzes the implications of considering the three dimensions of ability on the decision of attending four-year college. We find that, despite the high return associated with college attendance, individuals with low levels of cognitive and socio-emotional ability but high mechanical ability could expect higher wages by choosing not to attend a four-year college. These results highlight the importance of exploring alternative pathways to successful careers for individuals with a different profile of skills.

ESSAYS ON THE ECONOMICS OF ABILITY, EDUCATION AND LABOR MARKET OUTCOMES

by

María Fernanda Prada

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2014

Advisory Committee:
Professor Sergio Urzúa /Advisor
Professor Roger Betancourt
Professor Raymond Guiteras
Professor Soohyung Lee
Professor Joan Kahn

Acknowledgments

I am greatly indebted to Sergio Urzúa for his dedicated advising, encouragement and endless commitment. His work kindled my interest in the effects of abilities on labor market outcomes and his support empowered me to start and to persist in this project. I am greatful for his contributions to my dissertation and my academic training. I am thankful to Soohyung Lee for her invaluable guidance and to Raymond Guiteras for his helpful comments and continuous support. I also thank Roger Betancourt and Joan Kahn for agreeing to participate in the dissertation committee.

I am extremely grateful to John Ham, Miguel Sarzosa, Judy Hellerstein, Melissa Kearney, Lauren Deason and Giordano Palloni for the time they devoted and their contributions to this dissertation and my academic training.

I am deeply thankful to my family for their love, support, and sacrifices. Without them, this thesis would never have been written. This last word of acknowledgment I have saved for my loving, patient, encouraging and always positive husband Eduardo, whose faithful support will be always appreciated, and for Emilia whose existence has changed me forever.

Table of Contents

List of Tables

List of Figures

Chapter 1: **Introduction**

The analysis of the heterogeneity in worker ability and its economic implications have been a focus of a broad strand of research in labor economics. The multi-dimensional nature of ability implies that workers differ in both the level and the composition of their ability, which in turn represents differences in their productivity.

Over the last decades, several studies have demonstrated that both cognitive and socio-emotional dimensions of ability play an important role on market productivity as measured by wages, on the acquisition of skills and education, and on the choice of occupation. The prevalent result is that both dimensions of ability have a positive effect on outcomes. Higher levels of ability increase wages, the probability of progressing to higher levels of education and the probability of choosing jobs in high paying occupations.

However, there is no theoretical reason to expect that all dimensions affect outcomes in the same direction. In fact, some authors have shown the importance of another dimension of ability that is positively associated with wages, but implies different schooling, entrepreneurial, and occupational choices.[1]

In chapters 3 and 4 of this dissertation I study the role of mechanical ability

[1] Willis and Rosen (1979), Hartog and Sluis (2010), Yamaguchi (2012) and Boehm (2013)

as another dimension that, jointly with cognitive and socio-emotional dimensions, affects schooling decisions and labor market outcomes. I show that this dimension has a positive economic return and affects schooling decisions and occupational choices differently than other measures of ability.

This dissertation contributes to the literature in two major ways, both of which have ample productivity and policy implications. First, by expanding the range of empirically relevant dimensions of ability that I turn, enriches the current knowledge on the composition of human capital. Second, by presenting evidence to question the dichotomous paradigm of low and high ability individuals, in the context of the previously accepted symmetry of the impact of ability on important determinants of wages such as schooling decisions, occupational choices, and labor market productivity.

My analysis provides a better understanding of the dimensions of ability that are relevant to success in the labor market. This is important to define which dimensions we should foster as a society to increase productivity of labor force and also, to inform the debate on the conception of an educational system that develops and exploits the differences of individuals in terms of their abilities.

Chapter 2 introduces the concept of mechanical ability, the tests used to measure it and briefly compares this dimension with conventional measures of ability.

Chapter 3 contains the first essay where I develop the general framework used to understand the effects of multiple dimensions of ability on schooling choices, occupations and wages. I use an augmented Roy model that explicitly models two sequential selection processes and provides an estimation of counterfactual wages. I model the relationship between schooling, occupations, and wages simultaneously. Unlike other studies in the literature, I am able to decompose the total effect of initial unobserved abilities on wages into the components explained by schooling, occupation, and productivity on the job.

To account for the fact that workers sort into the occupations pursuing the tasks where their ability give them comparative advantage, I classify occupations as manual or abstract according to their core task requirements. This classification is inspired by the literature on tasks and job content Autor et al. (2003) as well as the skill-weights approach employed by Lazear (2003). My contribution here is twofold. First, I separate the source of identification for individual abilities and job characteristics by using tests scores to identify workers' abilities instead of infering them from the characteristics of the job. Second, I present a clasification as simple as the standard blue/white-collar that explains a larger fraction of the observed variance in wages.

Using data from the NLSY79, I find that all three dimensions of ability have multiple, heterogeneous, and independent roles. They influence the sorting of workers into schooling and occupations, and also directly affect wages, mainly by increasing productivity. Mechanical ability also increases wages but, unlike cognitive and socio-emotional ability, it is associated with low schooling levels and manual occupations. The productivity effect from mechanical ability is large enough to override the negative, indirect wage effects that work through schooling and occupational choice.

The results from the decomposition show that even if it is too late to change the schooling decisions or even the career path of individuals, interventions that increase ability can boost productivity and in consequence, wages of individuals late in their careers. In this context, the results from this dissertation inform the debate on the range of interventions that are relevant to increase productivity at different points in time.

Chapter four presents the second essay that was written in collaboration with Sergio Urzúa. In this essay we analyze the implications of considering a broader definition of ability in explaining the decision of attending four-year college. Using a simplified model that only contemplates schooling choice, we confirm the findings of the extended model in chapter 3 in the sense that all three dimensions have positive rewards on the labor market and mechanical ability is associated with low schooling levels. Our results suggest a new framework where individuals with low levels of cognitive and socio-emotional ability, may have high mechanical ability and greatly benefit from it. More precisely, we find that despite the high return associated with college attendance, these individuals could expect higher wages by choosing not to attend a four-year college. This conclusion is a direct result of the high returns to mechanical ability in jobs not requiring a four-year college degree which contrast with the negative returns to mechanical ability in jobs requiring it.

The results from our empirical model highlight the importance of moving beyond the “one-size-fits-all” college discourse and explore alternative pathways to successful careers for individuals with a different profile of skills. This message is particularly relevant in a nation where less than half of the students attempting to get a bachelor’s degree actually get one and where completion rates are below 20 percent for students who score low in standardized achievement tests during high school. Accepting the multidimensional nature of ability must be accompanied by the implementation of inclusive human capital development strategies with more than one pathway to success.

Chapter 2: **Mechanical Ability**

This chapter discusses conceptually mechanical ability, the tests used to measure this dimension of ability, and presents a comparison with conventional measures of ability.

2.1 Beyond Conventional Taxonomy

A large fraction of the literature on the effect of ability on schooling, labor market outcomes, and social behaviors has concentrated on cognitive skills: brain-based skills that are related to the mechanisms behind learning, remembering, problem-solving, and paying attention. In recent years, this literature has successfully incorporated socio-emotional abilities (e.g., persistence, grit, self-control, self-esteem) into the analysis. For example, Heckman et al. (2006) presents strong evidence of the importance of personality traits in explaining economic outcomes and a range of social behaviors. The same traits had already been linked to economic behavior by sociologists and psychologists (see, e.g. Bowles and Gintis, 1976; Edwards, 1976; Jencks, 1979; Wolfe and Johnson, 1995, among many others).

However, there might be other potential dimensions of ability determining, for example, human capital accumulation and labor market productivity. Indeed, common sense suggests that motor, manual dexterity, or even physical abilities may give an advantage to individuals in the labor market, specially if they are employed in certain occupations. I study a dimension of ability related to these aspects and label it mechanical ability. I borrow the name from the set of ability measures (test scores) available in the data, although I recognize that previous work has used a similar terminology.

But beyond its name, defining mechanical ability is a complex task. Cognitive and vocational psychologists as well as neuroscientists have utilized concepts such as mechanical aptitude, mechanical reasoning, and mechanical sense to describe this dimension.[1] Nevertheless, two distinctive components emerge from multiple definitions of mechanical ability. The first component, commonly named mechanical reasoning, is related to the ability to perceive and understand the movement or function of a mechanism either from interacting with it or by observing the mechanism. The second component is related to the ability to describe a mechanism that when, given some specified input, will produce a desired output (Blauvelt, 2006).

On the empirical side of this literature, the rising of the field of industrial psychology has fueled the interest in identifying the underlying traits leading to success in specific careers and occupations.[2]

On the other hand, the recent research on cognitive analysis, conducted by cognitive psychologists and neuroscientist, has focused on understanding how people

[1] See Blauvelt (2006) for a detailed literature review.

[2] Studies from vocational psychologists emerged early in the twentieth century Stenquist (1923), Cox (1928), Paterson et al. (1930). In particular, Cox (1928) and Paterson et al. (1930) were interested in finding a special mechanical intelligence which was separate from and complementary to Spearman's general intelligence quotient Spearman (1923).

reason mechanical devices and concepts. More specifically, this research has provided insights into how the brain acquires, processes, and uses information about mechanisms and machines.[3] This explains why most of the literature seeking to define mechanical ability focuses on the identification of rules used by the individuals to accomplish these tasks and to account for individual differences in performance.[4] Studies from neuroscientist concentrate in more specific abilities and the parts of the brain activated when performing different tasks. The main abilities identified by these types of studies relate directly to visual-motor integration and the visuospatial reasoning factors of spatial perception and spatial visualization (Hegarty et al., 1988; Carpenter and Just, 1989; Hegarty, 1992).

In economics, the few attempts trying to understand the role of mechanical abilities have examined its predictability power over schooling and labor market outcomes. Willis and Rosen (1979) included mechanical scores and manual dexterity tests in their study of college enrollment based on future labor market outcomes, obtaining that these dimensions reduce the probability of pursuing a college degree. My results are consistent with this unexplored finding, although they are not fully comparable given the differences in sources of information and empirical approaches betIen the two papers. Yamaguchi (2012) on the other hand, computes a measure of motor skills in his analysis of occupational choices throughout the life cycle. He finds that motor skills explains a large fraction of the observed wage variance and

[3] Most of the research from cognitive psychologists was produced during the 1980's Hegarty et al. (1988), Hegarty (1992), Carpenter and Just (1989), Heiser and Tversky (2002) to name a few.

[4] And in consequence to investigate the processes that distinguish people who score high or low in psychometric tests of mechanical ability.

also a large fraction of wage growth but only for high school dropouts.[5] In addition, Hartog and Sluis (2010) and Boehm (2013) use a measure of mechanical ability similar to the one analyzed below to study the characteristics of entrepreneurs, the sorting into middle skill occupations affected by polarization, respectively. I use it in chapter 3 to analyze early occupational choices.

The line of research started by Autor et al. (2003) has influenced these recent papers. In particular, the literature on task and skill content of jobs has provided a theoretical foundation for the analysis of the heterogeneity of worker's talent and the relationship with the variety of tasks required in the labor market. Mechanical ability can loosely be related with the type of skill needed to perform manual work that is intensively carried out by middle-education occupations.[6]

By analyzing the role of mechanical, cognitive and socio-emotional ability in the context of a schooling decision model with counter factual adult wages, I continue and extends the previous literature.

2.2 ASVAB: Technical Composites

The Armed Services Vocational Aptitude Battery (ASVAB) is a general test measuring knowledge and skills in the following areas: arithmetic reasoning, word knowledge, paragraph comprehension, mathematics knowledge, numerical operations, coding speed, general science, auto and shop information, electronics infor-

[5] It is important to note that the author does not take into account the endogeneity of the schooling decision and thus it is difficult to separate the effect through selection from the productivity effect.

[6] I present a more in depth discussion of this point in chapter 3.

mation, and mechanical comprehension.[7]

The literature has extensively analyzed the ASVAB, but typically focusing on the computation of the Armed Forces Qualification Test (AFQT). This test is used by the military services to determine basic qualification for enlistment, and its test score has been widely used as a measure of cognitive skills in economics (see, e.g. Cameron and Heckman, 1998, 2001; Ellwood and Kane, 2000; Heckman, 1995; Neal and Johnson, 1996; Heckman and Kautz, 2013, among many others).

To measure mechanical ability I use the following three sections of the ASVAB, commonly referred as the Technical Composites: the mechanical comprehension, auto and shop information, and electronics information sections. These sections are not used to compute the AFQT; instead, they are designed exclusively to compute the Military Occupational Specialty (MOS) scores.[8]

The questions from the *mechanical comprehension* section measure the ability to solve simple mechanics problems and understand basic mechanical principles, and represent one of the most widely used test measuring mechanical ability. They deal with pictures built around basic machinery such as pulleys, levers, gears, and Idges and ask to visualize how the objects would work together. People who understand mechanical devices can infer the principles of operation of an unfamiliar device from their knowledge of the device's components and their mechanical interactions (Carpenter and Just, 1989).

[7] The ASVAB is administered by the United States Military Entrance Processing Command and it is used to determine qualification for enlistment in the United States Armed Forces.

[8] The scores on these sections are used by the military to determine aptitude and eligibility for training in specific career fields within the military. Military career areas that require high scores on these three sections of the ASVAB include combat operations, general maintenance, mechanical maintenance, and surveillance and communications.

Moreover, the questions also cover topics such as how to measure the mass of an object, identify simple machines, and define words such as velocity, momentum, acceleration, and force. Some questions ask about the load carried by people or by support structures such as beams or bridges. For example, after showing a diagram with support structures, the question typically asks which one is the strongest or the Iakest, or which support in the diagram is bearing the lesser or greater part of the load. Many of the problems require basic mathematical skills such as knowledge on how to divide, work with decimals, and multiply two digit numbers.

The questions from the other two sections are similar to the mechanical section in that they require the ability to understand how objects work, but in the context of automotive and shop practices and electronics.

The *automotive and shop information section* measures technical knowledge, skills, and aptitude for automotive maintenance and repair and for wood and metal shop practices. The test covers the areas commonly included in most high school auto and shop courses, such as automotive components and requires an understanding of how the combination of several components work together to perform a specific function. It also includes questions on types of automotive and shop tools, procedures for troubleshooting and repair, properties of building materials, and building and construction procedures.

The *electronics information section* requires additional knowledge of the principles of electronics and electricity. For example, knowledge of electric current, circuits, how electronic systems works, electrical devices, tools, symbols, and materials is tested. Many of the topics covered in this section are probably covered in high school science classes.[9]

Although the questions ansIred by the respondents of the NLSY79 are not available, in Figure 4.1, I present sample questions obtained from the mechanical comprehension section. The two other sections are similar but they include topic specific terms and devices.[10]

The technical composites of the ASVAB have been proven to measure abilities and skills important to predict membership, training success, satisfaction, and job performance in the following career fields within the military: combat operations, general maintenance, mechanical maintenance, and surveillance and communications (Wise et al., 1992). Furthermore, according to Bishop (1988), the universe of skills and knowledge sampled by the mechanical comprehension, auto and shop information, and electronics subtests of the ASVAB roughly corresponds to the vocational fields of technical, trades and industry measured in occupational competency tests.[11] As a consequence, the three subtests of the ASVAB are interpreted as indicators of competence in these areas. All in all, the Technical Composites of the ASVAB should be viewed as measures of knowledge, trainability, and generic competence for a broad family of civilian jobs involving the operation, maintenance,

[9]An obvious concern for the identification strategy is the potential association betIen the *automotive and shop information* and *electronics information* sections and the material covered in specific classes during high school. This could potentially generate double causality betIen human capital accumulation and abilities. I follow Hansen et al. (2004) and deal with this potential source of bias by restricting the analysis to the youngest cohort of individuals in the sample as Ill as by controlling for the highest grade attended by the time of the test. I describe this strategy bellow. In addition, I analyze a small subsample of males for which I have high school transcript information, so I can confirm that they have not taken any elective course related to mechanical skills at the time of the tests. Results are qualitatively the same.

[10]I present a list of sample questions for the three sections in appendix A.

[11]Notable examples of occupation specific competency examinations are those developed by the National Occupational Competency Testing Institute and by the states of Ohio and New York to assess the performance of their high school vocational student. See Bishop (1988) for more detail.

and repair of complicated machinery and other technically oriented jobs (Bishop, 1988).

2.3 Measurement of Mechanical Ability in Perspective

In order to establish the relationship between the measure of mechanical ability and standard measures of ability, I show the correlation between the different tests. I also present the results from an Exploratory Factor Analysis that confirms the presence of one factor that is captured by the technical composites, but it is not captured by the other tests.

Table 4.1 shows the correlation matrix between the three technical composites of the ASVAB (Auto and shop information, mechanical comprehension, and electronics information), six tests used to compute AFQT (arithmetic reasoning, word knowledge, paragraph comprehension and math knowledge), the computed AFQT, and a composite measure of socio-emotional ability computed using Rosenberg Self-Esteem Scale and the Rotter Internal Locus of Control Scale. The three technical composites of the ASVAB are highly correlated with the scores in the questions used to compute AFQT, between 0.24 and 0.66, but present a low correlation with a standard measure of socio-emotional ability, between 0.18 and 0.21.

This is consistent with modern psychological theory which views ability as multidimensional with dimensions that are positively correlated with each other (Dickens, 2008). The positive correlation across abilities could be a manifestation of a general ability, sometimes referred to as the "Spearman g" or g-factor Spearman

(1904), or could be the result of overlap in the knowledge required to answer the different tests.[12]

Further analysis of the correlation among the variables used to create AFQT and the technical composites highlights the presence of two different components. The results form an Exploratory Factor Analysis (EFA) on the nine subsections of the ASVAB (the three technical composites plus the four set of questions used to create the AFQT) confirm that at least two factors are needed to explain the correlation among the scores in the nine questions.[13]

All the loadings corresponding to the first factor are positive and statistically significant, they range between 0.62 and 0.83. In contrast, the loadings for the second factor differ between the questions used to compute the mechanical ability measure and the questions used to compute AFQT. More specifically, for the three tests used to construct the mechanical measure the loadings are high and statistically significant, they range between 0.31 and 0.48 but for the rest of the tests, the loadings are close to zero.[14] Panel a) in Figure 4.2 presents the original estimated loadings.

The results from the EFA suggest a structure where the first factor is important to linearly reconstruct all questions but the second factor is only relevant for the

[12]More specifically, it could be explained by the fact that all the questions in the three composites of the ASVAB require a certain degree of reading or verbal comprehension or that many of the problems require basic mathematics skills.

[13]In addition, the factor analysis assuming orthogonal factors and allowing for some unique components in the equation keeps the four first factors, because the default criteria is to keep all the factors with positive eigenvalues. The eigenvalue for the first factor is 4.75 and 0.80, 0.22 and 0.17 for the next three factors. The first two factors account for all the shared variance, 85 percent the first and 15 percent the second, so I focus only on them.

[14]Numerical Operations is an exception because the laoding for the second factor is highly negative (-0.38). The magnitude of the loading is critical because any factor loading with an absolute value of .30 or greater is considered significant (Diekhoff, 1992; Sheskin, 2004, among others).

three technical composites of the ASVAB. Figure 4.2 presents the estimated loadings for each factor, i.e., the estimated coefficients associated with each factor. The suggested structure persists also after several forms of rotation.[15]

In this context, the first factor is capturing all the common information that is expressed by the high positive correlation among the tests and the second factor captures the additional component that makes the three tests used to measure mechanical ability different from the AFQT.

We assume that the first factor, shared by all components of the ASVAB, is measuring cognitive ability. This factor affects the three technical composites of the ASVAB because several questions require a certain degree of reading or verbal comprehension and basic mathematics skills associated with cognitive ability. The second factor, which is only present for the technical composites, may be related to mechanical ability. The part of ability that is related to understanding how things work but it is not captured by the AFQT. I incorporate this ideas in the empirical model.

If one wants to describe a trilogy of abilities that are rewarded in the labor market I can be said that cognitive abilities capture *conceptual and thinking skills*, while socio-emotional/socio-emotional skills capture human relations skills ,*people skills* and mechanical would be more related to technical skills *how-to-do-it skills*.

[15]Rotation is important because of the indeterminacy of the factor solution in the exploratory factor analysis. In panel b) of Figure 4.2 I present the loadings after a rotation made to maximize the variance of the squared loadings between variables (simplicity within factors).

Chapter 3: **Beyond Smart and Sociable: Re-thinking the Role of Abilities on Occupational Choices and Wages**

3.1 Introduction

The assignment of workers to the tasks where they can be most productive is a fundamental issue in economics. Starting with the seminal work of Roy (1951) on self-selection, numerous studies have analyzed the sorting of heterogeneous workers into the occupations where they have comparative advantage. An essential contribution of the Roy model is the formalization of the notion that there are multiple dimensions of ability and that these dimensions differ in terms of how relevant they are for distinct occupations.

In this chapter, I study the effects of multiple dimensions of ability on early occupational choices and productivity, measured in wages. I concentrate on the stock of abilities owned before choosing the final level of schooling and also before entering into the labor market. My analysis has three main contributions to this literature.

First, I explore the implications of the multidimensional nature of ability by extending the traditional cognitive-noncognitive framework to include mechanical ability. Mechanical ability is strongly associated with productivity in a particular class of occupations. It is also an important predictor of wages but, unlike cognitive and socio-emotional abilities, it has different implications in terms of schooling and occupational choice.

Second, I classify occupations according to their core task requirements. This allows me to study occupational choices in terms of the association between worker's ability and the activities performed at the job. Following the literature on tasks and job content (see for example Autor et al., 2003) and also the skill-w8ths approach employed by Lazear (2003), I use data from the O*NET to classify occupations as manual or abstract. Unlike the common approach that uses characteristics of the job to infer workers' abilities, in this paper workers' talents are identified from individual tests.[1] As a consequence, individual abilities are not themselves directly associated with occupational categories.

Finally, I model the relationship between schooling, occupations, and wages simultaneously. This enables me to identify all of the channels through which abilities affect outcomes. The existing literature analyzes the effect of abilities on each of these decisions separately.[2] In contrast, I am able to decompose the total effect of abilities into the components explained by schooling, occupation, and productivity.

[1] Autor et al. (2003); Ingram and Neumann (2006); Autor and Dorn (2009); Poletaev and Robinson (2008); P. and Blu (2010); Yamaguchi (2012, among others). Autor and Michael (2013)

[2] Willis and Rosen (1979), among others concentrate on schooling decisions while others concentrate on occupational self-selection abstracting from the endogeneity of schoolingWillis (1986), Rubinstein and Weiss (2006), Yamaguchi (2012), Gibbons et al. (2005). Heckman et al. (2006) incorporate both schooling and occupational decisions, but they are not interrelated. Notable exceptions are Keane and Wolpin (1997), Lee (2005) and Sullivan (2010)that include all components but the source of differences in unobserved ability cannot be identified, only partially characterized with ex-post realizations.

A recent and growing literature on cognitive and socio-emotional abilities has concentrated on exploring worker heterogeneity and its consequences for schooling, labor market outcomes, and other behaviors.[3]

The prevalent result is that both cognitive and socio-emotional dimensions of ability have a positive effect on outcomes. For example, both increase the probability of progressing to higher levels of education, increase the probability of choosing jobs in high paying occupations, increase wages, etc.

But, there is no reason to expect that all dimensions affect outcomes in the same direction. In fact, Willis and Rosen (1979), Hartog and Sluis (2010), Yamaguchi (2012) and Boehm (2013) among others have shown the importance of another dimension of ability that is positively associated with wages, but implies different schooling, entrepreneurial, and occupational choices.[4]

Furthermore, abilities can affect multiple outcomes without necessarily being direct determinants of occupational choices and market wages. Instead, they might influence outcomes by changing preferences, endowments, the efficiency of human-capital production or school performance.[5] For example, abilities might indirectly

[3]See Bowles and Gintis (1976); Herrnstein and Murray (1994); Murnane et al. (1995); Neal and Johnson (1996); Duncan and Dunifon (1998)Cawley et al. (2001); Carneiro and Heckman (2003); Heckman et al. (2006); Cunha et al. (2006); Duckworth et al. (2007); Urzua (2008); Borghans et al. (2008); Duckworth and Urzua (2009); Conti et al. (2010); Ferguson et al. (2011), Hartog and Sluis (2010); Tambunlerchai (2011); Sarzosa and Urzua (2013)and many others.

[4]Willis and Rosen (1979) analyze the decision of going to college; the former using mechanical scores and manual dexterity as indicators of ability and the latter using the technical composites of the ASVAB to estimate "mechanical" ability. Both studies find that this dimension of ability predicts lower levels of schooling and analyze its effect on wages but do not consider the role of occupation in explaining the observed differences on wages. On the other hand, Yamaguchi (2012) analyzes occupational choices throughout the life cycle and Hartog (2001) studies the choice of being an entrepreneur. Both find that this dimension of ability predicts the choice of occupations associated with lower wages but the economic returns on those occupations is very high. Neither study takes into account the endogeneity of the schooling decision and thus it is difficult to separate the effect through selection from the productivity effect.

[5]See Cunha and Heckman (2007); Cunha et al. (2006, 2010).

impact occupational choice through their effect on schooling by determining the number and type of occupations available to the worker. Cognitive, socio-emotional, and mechanical skills might also raise the productivity of workers in different occupations and thereby directly affect wages.

In this context, the objective of this paper is to understand the main channels through which the three dimensions of ability affect occupational choices and wages. How important are worker's pre-labor market abilities on early occupational choices? Do the different dimensions of ability retain explanatory power after accounting for their influence on schooling? What portion of the total effect of ability on wages is explained by a direct productivity effect? Finally, does mechanical ability help to understand behaviors and decisions that could not be explained using the cognitive/socio-emotional framework?

To answer these questions, I use an augmented Roy model with a factor structure that explicitly models two sequential selection processes. This model closely follows the model presented in Heckman et al. (2006) and Urzua (2008). Workers first decide their level of schooling, taking into consideration their abilities. Then, they workers select into occupation based on their abilities and their previous schooling choices. I use observed measures of abilities (test scores) to identify the distribution of unobserved cognitive, socio-emotional and mechanical abilities. For the empirical analysis, I use data of young white males from the NLSY79.

I find that all three abilities have multiple, heterogeneous, and independent roles. They determine the sorting of workers into schooling and occupations. Cognitive and socio-emotional ability are associated with high levels of schooling and selection into abstract occupations. Mechanical ability, on the other hand is associated with low schooling levels and manual occupations.

Each component of ability directly affects the choice of occupation according to the main tasks required in the job. In addition, a sizable fraction of the total effect of pre-labor market abilities on occupational choice is driven by the indirect effect through schooling: nearly 40 percent for cognitive and mechanical ability and 25 percent for socio-emotional. This indirect effect presumably captures how different schooling levels alter the choice set of occupations available to workers.

All three dimensions of ability increase average wages. A one standard deviation increase in cognitive, socio-emotional, and mechanical skills lead to a 12 percent, 6 percent, and 2.7 percent wage increase, respectively. Moreover, all three dimensions of ability have a sizable productivity effect. For cognitive skills, 33 percent of the total effect is explained by increased productivity and 35 percent by increased schooling attainment. For both socio-emotional and mechanical skills the majority of the total effect can be attributed to the direct productivity channel.

In contrast to cognitive and socio-emotional ability, mechanical ability is associated with lower schooling attainment and a different profile of occupational choices. In addition, the direct, productivity effect of mechanical ability is considerably higher than it is for either cognitive or socio-emotional ability. In fact, the positive impact of mechanical ability on productivity (wages) is large enough to entirely offset the negative, indirect impact which results from the lower implied schooling level and the choice of manual occupations.

This document contributes to the literature on heterogeneous human capital and occupational choice by analyzing the role of specific components of the vector of initial endowments (pre labor-market abilities) instead of a generic composite of initial endowments that are occupation-specific.[6] In particular, I explore the heterogeneity of the unobserved component of initial ability, which is an important determinant of ex-post differences in wages and lifetime welfare.[7]

In addition, I relax the restriction for abilities to be normally distributed, as is generally assumed in the literature.[8] I argue that unobserved abilities are not normally distributed and also that the estimated distributions of abilities imply a different sorting into schooling and occupations than would be implied by observed test scores.

The remainder of the paper is organized as follows. Section 2 explains the procedure followed to classify occupations into manual-abstract categories according to the task requirements. Section 3 describes the NLSY79 data used to estimate the model and highlights the overall patterns in the data. Section 4 presents the model and Section 5 the estimation strategy. The estimation results including the model fit are presented in Section 6. Section 7 discusses the decomposition of the total effect of each type of ability into its distinct components. Conclusions are presented in Section 8.

[6] The great majority of the literature concentrates on occupation specific skills as in Keane and Wolpin (1997); Rubinstein and Weiss (2006); Kambourov and Manovskii (2009) Sullivan (2010), Antonovics and Golan (2012) among others.

[7] For example,Keane and Wolpin (1997) find that 90 percent of the total variance in expected lifetime utility is explained by differences in skill endowments. However it is not possible to determine the sources of the differences in initial endowments. Sullivan (2010) and Yamaguchi (2012) finds that skill endowments explain more than 70% of the observed variance in log wages but again endowments. Although they find that the importance of endowments fades with time, after 20 of experience initial endowments still explain an important percent of the variance in wages, close to 35 percent.

[8] Willis and Rosen 1979; Yamaguchi 2012, and many others

3.2 Using Task Content to Classify Occupations

As previously stated, one of the contributions of this paper is to analyze occupational choices in terms of how people skills relate to the activities predominantly required in each occupation. To this end, I classify occupations into two categories according to the core task requirements of jobs instead of other criteria such as responsibilities, people in charge, industry, education, etc. In this section I describe the classification procedure and compare it with the standard white-collar/ blue-collar classification.

I assume that tasks are broadly categorized into either abstract tasks or manual tasks. This is in the spirit of the original classification proposed by Autor et al. (2003) but without the emphasis on routine vs non-routine tasks.[9] As in Acemoglu and Autor (2011), abstract tasks are activities that require problem-solving, intuition, persuasion, creativity, and in-person interactions. Manual tasks are activities that require the use of the hands, and the physical body (musculoskeletal system) to perform work, including the use and manipulation of external objects such as tools, machinery, etc.

To do the classification, I use information from the Occupational Information Network (O*NET), the successor of the U.S. Department of Labor's Dictionary

[9] Autor et al. (2003) consider five task groups: Non-routine analytic, Non-routine interactive, routine cognitive, routine manual and non-routine manual. This classification was made to separate tasks according to their relationship with computers in order to understand the role of technological change in the labor market. This classification does not meet the purposes of this paper. My definition of manual tasks is closer in spirit to their manual non-routine. After Autor et al. (2003) a growing number of papers have adopted the so called "task-approach". Some use the DOT and others O*NET, depending the specific purposes. Autor et al. (2006)(2008), Goos and Manning (2007), Peri and Sparber (2008), Goos, Manning and Salomons (2010), Autor and Dorn (2009), Borghans et al. (2007), Acemoglu and Autor (2011),Yamaguchi (2012),Firpo et al. (2011), Autor and Michael (2013)among many others. Most of the authors use the categories of Acemoglu and Autor (2011) with variations depending on the specific objectives of their analysis.

of Occupational Titles (DOT) to create task measures and then impute them to workers in the NLSY79 according to their occupation.[10] This imputation allows me to convert several hundred occupational titles found in conventional occupational classifications into just two task dimensions.

The O*NET database contains detailed information on over 900 occupations. For each occupation, it provides a list of required work activities divided in four main categories: information input, mental processes, work output, and interacting with others. Each activity has two scores, one associated with its "importance" for the job and the other associated with its "level" (degree of complexity). I use only the "importance" score for because both are highly correlated so they contain almost the same information but the importance score is easier to interpret.[11]

One disadvantage of the wealth of information of the O*NET is that it is not obvious as to how to create a measure that best represents a given task construct. To overcome this, I do a Principal Component Analysis (PCA) on all of the item/s in each subcategory. Then, I select the items with the highest coefficient (loading) on the first component, the component that summarizes most of the common information among all the items in the subcategory.[12]

This process uses the following 6 items to create the measure of abstract task

[10] I use ON*NET Version 17.0. July 2012 release. http://www.onetcenter.org/database.html.

[11] The correlation between "importance" and "level" scores is close to one, so there is little information added by using both.In addition, the "importance" question is more standard and easier to interpret. It is based on a scale from 1 to 5, monotonically increasing in the importance of the activity for the job; while, the "level" questions ask respondents to choose one position on a 1 to 7 scale. Although examples for positions 2, 4, and 6 are provided as a benchmark, the examples are not always informative and it is not straightforward to position an activity in between two example. In fact, in pilot studies and subsequent evaluations, occupational analysts found it difficult to interpret level ratings. http://www.onetcenter.org/dl_files/AOSkills_ProcUpdate.pdf. Figure 3.1 present an example of the questionnaire.

[12] If two items inside a subcategory have extremely similar loadings, I use both.

complexity: three related to analytical skills; 1) Analyzing data/information, 2) Thinking creatively, 3) Organizing, planning and prioritizing work; and three related to interpersonal skills; 4) Establishing and maintaining personal relationships, 5) Guiding, directing, and motivating subordinates, and 6) Coaching and developing others.

For the manual task complexity measure, the following six items were used: 1) Controlling machines and processes , 2) Handling and moving objects, 3) Repairing mechanical/ electrical equipment, 4) Time spent using hands to handle, control, or feel objects, tools, or controls, 5) Manual dexterity, and 6) Visualization.

Each scale is then standardized to have mean zero and standard deviation one. The composite task measures used are equal to the summation of their respective constituent scales, then standardized to have mean zero and standard deviation one. In order to merge the composite task measures with the NLSY79 data, they are collapsed to the Census 1990 occupational code level using the Census 1990 labor supply weights, and then collapsed to the 396 consistent occupations as detailed in Autor and Dorn (2009).[13]

Finally, for each occupation I compare the ranking of the manual and abstract composite task measures with the distribution in the population. An occupation is classified as abstract if the position of the abstract task measure in the distribution is higher than the position of the manual measure.

[13]Another alternative is to collapse the original categories proposed by Autor et al. (2003) into the two that are relevant for this study, manual and abstract. Although, the classification of occupation between abstract and manual does not change significantly, my approach is more neutral and relies less on the routine/non-routine differences between occupations. In addition, the dummy created with the classification utilized in the paper explain a greater percentage of the log wage variance. See appendix for details.

The proposed classification is as simple as the standard white/blue collar but provides a systematic way of classifying occupations according to the main job requirements. Figure 3.2 presents the comparison of the manual and abstract composite task measures for blue and white-collar workers in the sample. For most of the occupations classified as blue-collar the manual task measure is higher than the abstract task. The same is true to a lesser extent for white collar occupations.

One advantage of this definition is that it classifies at least three types of occupations more appropriately: 1) services that are classified as blue collar but are similar in tasks to white-collar jobs, such as high rank police chief, detectives, etc; 2) Technicians whose characteristics may resemble other white collar workers but spend most of their time working with machines as their blue-collar counterparts; and 3) factory operatives, working in industrial laboratories who are classified as blue-collar but performing activities and tasks similar to technicians.

This classification is flexible enough to capture large variations across three-digit census occupations that are generally grouped into the same one-digit occupation. Figure 3.3 presents one example. It compares the centiles of manual and abstract measures for different occupations. On average, the occupations typically classified as white-collar (Professional, managers, sales and cleric) also would be classified as abstract occupations (see panel A). However, in analyzing the measure at a greater level of detail it is evident that some white-collar occupations that have high manual requirements, as in the case of technicians, are classified as manual. For more details on the comparison between the two classification of occupation and the implication on outcomes see Appendix 4.

3.3 Data: National Longitudinal Survey of Youth 1979

This section describes the data utilized to estimate the model and some descriptive statistics.

The National Longitudinal Survey of Youth (NLSY) is a panel data set of 12,686 individuals born between 1957 and 1964.[14] This survey is designed to represent the population of youth aged 14 to 21 as of December 31 of 1978, and residing in the United States on January 1, 1979. It consists of both a nationally representative cross-section sample and a set of supplemental samples designed to over-sample civilian blacks, civilian Hispanics, economically disadvantaged Non-Black/ Non-Hispanic youths and individuals in the military. Data were collected on an annual basis from 1979 to 1994 and biannually until present day.

I use the sample of white males at age 26 who were not attending school at the time of the survey and who had not yet graduated from high school at the time the tests used to measure ability were collected (Survey of 1979 and the summer 1980). I exclude females and non-white males from the analysis to concentrate on the effect of abilities on schooling and labor market outcomes while abstracting from other important forces such as discrimination and gender preferences. In addition, by age 26 nearly 87 percent of the sample has reached their maximum level of education, so the analysis here concentrates on final schooling choices, rather than intermediate states.

From the original sample of 12,686 individuals, 11,406 are civilian and 6,111

[14] 5,579 males-49 percent of total surveyed individuals.

belong to the cross-sectional sample. Nearly 49 percent of that sample are males, 2,439 individuals are observed at least once at age 26, or between 25-27 for those not observed at 26. I exclude 540 individuals who had already completed high school by the time the ASVAB test was conducted (survey date in 1979 for socio-emotional tests and Summer 1980 for ASVAB test). This is relevant because the schooling margin I analyze is completing some college versus completing high school or not completing high school. Test scores are captured before the final decision on schooling is made and before any labor market experience. When I also exclude individuals attending school at the time the survey was conducted, the sample is reduced to 1,655 individuals. Table 3.1 presents a description of the variables used for the final sample of individuals with available information on all the variables of interest to compute the schooling choice decision equation. Wages and occupation categories have few data because they depend on the participation of the individual and the availability of information on the occupation category.

I analyze one schooling choice: pursuing some education beyond high school or not. The variable used to measure this choice is the highest degree completed by the age of 25. The labor market outcomes I analyze are the occupation and log of hourly wages at the current or most recent job (CPS job).

For the cognitive and mechanical measures I rely on the Armed Services Vocational Aptitude Battery (ASVAB) that was conducted in the summer and fall of 1980.[15] This test was administrated to over 90 percent of the members of the NLSY

[15]This questions are used to compute the AFQT that is used by the military services for enlistment screening and job assignment within the military.

panel (individuals were between 15 and 23 years old at the time of the test).[16] The test is made up of a battery of 10 questions measuring knowledge and skills in the following areas: arithmetic reasoning, word knowledge, paragraph comprehension, numerical operations, coding speed, mathematics knowledge, general science, auto and shop information, mechanical comprehension and electronics information. The first 6 are used as measures of cognitive ability while the last 3 are measures of mechanical ability.

For measures of socio-emotional ability I use two tests: the Rotter Locus of Control Scale and the Rosenberg Self-Esteem Scale. The Rotter Locus of Control Scale measures the degree of control individuals feel they possess over their life. In 1979 the NLSY collected a total of four items selected from the 23-item forced choice questionnaire adapted from the 60-item Rotter Adult I-E scale developed by Rotter (1966).[17] The Rosenberg Self-Esteem Scale, which is based on 10 questions, measures self-esteem: the degree of approval or disapproval towards oneself Rosenberg (1965). The scale is short, widely used, and has accumulated evidence of validity and reliability.[18]

---

[16]As already pointed out I only use individuals who were not finished high school by the time of the test.

[17]"This scale was designed to measure 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 (that is, chance, fate, luck) controls their lives (external control). The scale is scored in the external direction-the higher the score, the more external the individual" Extracted from http://www.nlsinfo.org/nlsy79/docs/79html/79text/attitude.htm

[18]It contains 10 statements of self-approval and disapproval with which respondents are asked to strongly agree, agree, disagree, or strongly disagree. The scale has proved highly internally consistent, with reliability coefficients that range from .87 Menaghan (1990) to .94 Strocchia-Rivera (1988), depending on the nature of the NLSY79 sample selected. Ibid.

3.4 Model: Augmented Roy Model with Factor Structure

The model presented in this section deals with two of the main problems that arise when computing the effect of latent, initial abilities on occupation and wages: the endogeneity of both schooling and occupational choices and the fact that test scores are just proxies for abilities and they are influenced by schooling, age, and family background variables.

In the model, individuals are endowed with a three-dimensional vector of ability. These dimensions of ability jointly determine the schooling choices they make . Also, they are synthesized to perform the tasks involved in any occupation, but each occupation rewards each dimension of ability differently. As a consequence, the returns to each dimension vary by occupation and schooling.[19]

The strategy pursued in this paper is based on a model that integrates schooling decisions, occupational choices and wages. The model proposed closely follows the models presented in Heckman et al. (2006) and Urzua (2008) where a vector of low dimensional factors, in this case cognitive, mechanical, and socio-emotional abilities, is used to generate the distribution of potential outcomes. These latent abilities generate measured cognitive, socio-emotional, and mechanical scores and the rest of the outcomes analyzed. Conditioning on observables, these factors account for all of the dependence across choices and outcomes.

The theoretical model does not consider the exact timing of the decisions. However, the occupational choice model is assumed to differ by schooling level and

[19] In the standard Roy-Model setup skills are used in one occupation but not in the others.

this explains the sequential nature of the model. Agents choose their maximum level of schooling before the age 26 given the information they have at the time. Individuals then decide the type of occupation where they will work, after considering their previous schooling decisions. I employ this sequence of decisions to allow for schooling levels to restrict the type and number of occupations available to the worker. As a consequence, the choice between manual and abstract occupations is not the same at all education levels. In addition, I assume each individual receives an idiosyncratic shock between the time deciding on a schooling level and selecting an occupation. Finally, the schooling choice model is evaluated at the age 26.

I assume that latent abilities are unobserved by the econometrician but the individual has full information about his/her abilities, as well as knowledge of how they affect the potential earnings in each education-occupation cell. The agent compares the potential outcomes across each feasible choice and chooses the alternative that yields the highest payoff.

The structure of the model is described as follows:

$$Y = X\beta + u^Y$$

$$T = Q\Gamma + u^T$$

where Y is the vector of decisions and outcomes of interest (schooling decisions, occupational choices, accumulated experience in manual and abstract tasks, and wages), X includes a set of observable variables that explain outcomes (geographic and cohort controls) and u^Y is the error term. T is a vector of observed test scores,

Q includes a set of observable variables that explain test scores including family background characteristics and education at the time of the test, and u^T is the error term.

The error term is composed of three factors representing latent abilities and idiosincratic shocks, so the model can be rewritten as:

$$Y = X\beta + \lambda'\theta + e^Y$$

$$T = Q\Gamma + \gamma'\theta + e^T$$

Where θ is the vector of latent abilities $\theta = \{\theta_C, \theta_M, \theta_S\}$, λ and γ the vectors of returns to these abilities and $e^Y$and e^T are iid idiosyncratic shocks. I assume that the level of individual abilities is the result of some combination of innate ability, the quality of the environment provided by her parents, and her efforts and interventions before taking the tests. I also assume that the individuals have perfect information about their own abilities and that they are fixed by the time the individual makes her choices.

The vector of decisions and outcomes, Y, includes the schooling decision D, the choice of abstract occupation over manual occupations for each level of schooling, D_{s0} and D_{s1} and the potential wage for each combination of schooling and occupation: high school or less-manual occupation, high school or less-abstract occupation, some college or more-manual occupation, some college or more-abstract occupation.

$$
Y = \begin{bmatrix} D \\ D_0 \\ D_1 \\ Ln(W) \end{bmatrix} = \begin{bmatrix} Pr(D=1) = X_D\beta + \lambda'_D\theta + e \\ Pr(D_0=1) = X_0\beta_{D_0} + \lambda'_{D_0}\theta + e_{D_0} \\ Pr(D_1=1) = X_1\beta_{D_1} + \lambda'_{D_1}\theta + e_{D_1} \\ \begin{bmatrix} \ln w_{00} = X\beta_{00} + \lambda'_{00}\theta + e_{00} & \text{if } D=0 \ \&\ D_0=0 \\ \ln w_{01} = X\beta_{01} + \lambda'_{01}\theta + e_{01} & \text{if } D=0 \ \&\ D_0=1 \\ \ln w_{10} = X\beta_{10} + \lambda'_{10}\theta + e_{10} & \text{if } D=1 \ \&\ D_1=0 \\ \ln w_{11} = X\beta_{11} + \lambda'_{11}\theta + e_{11} & \text{if } D=1 \ \&\ D_1=1 \end{bmatrix} \end{bmatrix}
$$

The vector of test scores includes by the vector of cognitive, mechanical and socio-emotional tests, C, M and N, respectively.

$$
T = \begin{bmatrix} C = Q_c\Gamma_c + \gamma'_c\theta + e_c \\ M = Q_m\Gamma_m + \gamma'_m\theta + e_m \\ S = Q_s\Gamma_s + \gamma'_s\theta + e_s \end{bmatrix}
$$

Each of the components of the model will be presented in a separate subsection. The model estimated uses 2 schooling levels (high school or less versus some college or more), 3 factors (the three dimensions of ability), 6 cognitive tests, 3 tests on mechanical ability, and 2 tests on socio-emotional abilities.

3.4.1 Model of Schooling Choice

The latent utility of getting education is given by:

$$D = \mathbf{1}[I_i > 0]$$

$$I_i = X_{D,i}\beta + \lambda_D^c\theta_{c,i} + \lambda_D^m\theta_{m,i} + \lambda_D^s\theta_{s,i} + e_i \text{ for } i = 1, ...N$$

$$e_i \sim N(0, 1)$$

where $X_{D,i}$ is a matrix of observed variables that affect schooling, β is the vector of coefficients. $\widehat{\theta} = [\theta_{c,i}, \theta_{m,i}, \theta_{s,i}]$ is the vector of latent abilities where subscript c is used to denote the cognitive ability, subscript m denotes mechanical ability and subscript s denotes socio-emotional ability. $\lambda_D^c, \lambda_D^m, \lambda_D^s$, the vectors of returns to these abilities. These coefficients are referred in the literature as the factor loadings. e_i is the error component that is assumed to be independent of X_D, θ and following a standard normal distribution. Then D denotes a binary variable that takes the value of 1 if the individual chooses to attend a 4-year college and 0 otherwise.[20]

Conditional on X_D and θ the equations produce a standard discrete choice model with a factor structure. Furthermore, given the set of assumptions exposed, this can be interpreted as the standard probit model.

[20]Through all the exposition the indicator function will be used, $\mathbf{1}[]$ this function takes a value of one if the condition inside the parentheses is satisfied.

3.4.2 Model of Occupational Choice

The latent utility of working in an abstract occupation, conditional on the level of schooling attained is given by:

$$D_O = \mathbf{1}[I_{O,i} > 0]$$

$$I_{O,i} = X_{O,i}\beta_{D_O} + \lambda^c_{D_O}\theta_{c,i} + \lambda^m_{D_O}\theta_{m,i} + \lambda^s_{D_O}\theta_{s,i} + e_{O,i}$$

$$\text{for } i = 1, ...N \text{ and } O = 0, 1.$$

$$e_{O,i} \sim N(0, 1)$$

where O is an indicator for the final schooling level, $X_{O,i}$ is a matrix of observed variables that affect occupational choice given schooling level O, β_{D_O} is the vector of coefficients. $\widehat{\theta} = [\theta_{c,i}, \theta_{m,i}, \theta_{s,i}]$ is the vector of latent abilities and $\lambda^c_{D_O}, \lambda^m_{D_O}, \lambda^s_{D_O}$ the respective factor loadings. $e_{O,i}$ is the error component that is assumed to be independent of X_O, $\widehat{\theta}$ and following a standard normal distribution. Then D_O denotes a binary variable that takes the value of 1 if the individual with education O chooses to work in an abstract occupation and 0 otherwise.

Conditional on X_O and $\widehat{\theta}$ the equations produce a standard discrete choice model with a factor structure. Furthermore, given the set of assumptions exposed, this can be interpreted as the standard probit model.

3.4.3 Model of Hourly Wages

Analogously, the model of earnings can be expressed as a linear function of X_i and θ in the following way:

$$\ln w_{d,o,i} = X_{w,i}\beta_{d,o} + \lambda^c_{d,o}\theta_{c,i} + \lambda^m_{d,o}\theta_{m,i} + \lambda^s_{d,o}\theta_{s,i} + e_{d,o,i}$$

$$e_{d,o,i} \sim N(0, \sigma_{d,o})$$

for $d = \{0, 1\}$ and $o = \{0, 1\}$. where d is the indicatior for schooling level and o the indicator for occupation as before the value of zero indicates a manual occupation and the value of one an abstract occupation.

3.4.4 Model of Test Scores: Measurement System

Motivated for the findings of the Exploratory Factor Analysis performed in Section 3 the model of test scores allow each measurement to be a function of the corresponding latent ability. For the mechanical tests we allow them to be a function of both cognitive and mechanical latent factors.

In this context, the model for the cognitive measure C_j is:

$$C_{j,i} = X_{C_j,i}\beta_{C_j} + \lambda^c_{C_j}\theta_{c,i} + e_{C_j,i}$$

for $j = \{1, ..., 6\}$.

The model for the mechanical measure M_l is:

$$M_{k,i} = X_{M_k,i}\beta_{M_k} + \lambda^c_{M_k}\theta_{c,i} + \lambda^m_{M_k}\theta_{m,i} + e_{M_k,i}$$

for $k = \{1, ..., 3\}$.

And the model for the socio-emotional measure S_l is:

$$S_{l,i} = X_{S_l,i}\beta_{S_l} + \lambda^s_{S_l}\theta_{s,i} + e_{S_l,i}$$

for $l = \{1, 2\}$.

Finally, all error terms $\{e_i, e_{w,D,i}, e_{C_1,i}, ..., e_{C_6,i}, e_{M_1,i}, ..., e_{M_3,i}, e_{S_1,i}, e_{S_2,i}\}$ for $D = \{0, 1\}$, $j = \{1, ..., 6\}$, $k = \{1, ..., 3\}$ are mutually independent, independent of the factors and independent of all observable characteristics. This independence is essential to the model since it implies that all the correlation in observed choices and measurements is captured by latent unobserved factors.

3.4.5 Latent Factors

The observed level of these latent factors may be the result of some combination of inherited ability, the quality of the family environment in which individuals were raised, cultural differences, etc. These factors are assumed to be fixed by the time the individual is choosing the level of education and thus, by the time the labor and behavioral outcomes considered in this document are determined. In addition, the factors are assumed to be known by the individual but unknown to the researcher.

Following standard conventions it is assumed that cognitive and mechanical factors are independent to the Socio-emotional factor while cognitive and mechanical can be correlated.

A mixture of normals is used to model the distribution of the latent abilities. This distribution is selected because as Ferguson (1983) proved, a mixture of normals can approximate any distribution and we want to impose the minimum number of restrictions on the distribution of these unobserved components.

In this case, we use mixtures of two normal distributions (i.e., $K = J = L = 2$) and assume $E[\theta_c] = E[\theta_m] = E[\theta_s] = 0$. Finally, we impose $(\theta_c, \theta_m) \perp \theta_s$. For more details on this and the identification strategy refer to Appendix B.

3.5 Estimation

This section contains a brief explanation of the estimation strategy and presents the likelihood function associated with the estimation of the model.

Let $T_i = \{C_{1,i}, ..., C_{6,i}, M_{1,i}, ..., M_{3,i}, S_{1,i}, S_{2,i}\}$, be the vector of test scores for individual i. Let $\theta = [\theta^c, \theta^m, \theta^s]$ be the vector of the latent factors and δ the vector of all the parameters of the model

$$L(\delta|X, Q) = \prod_{i=1}^{N} f(D_i, D_{d,i} \ln w_{d,o,i}, T_i | X_{D,i}, X_{O,i}, X_w, Q_i)$$

Given that conditional on unobserved endowments, all the errors are mutually independent. Similar to previous papers Heckman et al. (2006); Urzua (2008) this can also be expressed as:

$$L(\delta|X,Q) = \prod_{i=1}^{N} \int_{\Theta} f(D_i, D_{O,i} \ln w_{d,o,i}, T_i | X_{D,i}, X_{O,i}, X_w, Q_i, \theta) dF(\theta)$$

where

$$f(D_i, D_{O,i} \ln w_{d,o,i}, T_i | X_{D,i}, X_{O,i}, X_w, Q_i, \theta) = f(D_i, D_{O,i} \ln w_{d,o,i}, | X_{D,i}, X_{O,i}, X_w, \theta) f(T_i | Q_i, \theta)$$

The model is estimated using MCMC techniques. The use of Bayesian methods in this paper is merely computational to avoid the computation of a high order integral. In consequence, the interest is primarily on the mean of the posterior distribution. Thus, it is viewed from a classical perspective and interpreted as an estimator that has the same asymptotic sampling distribution as the maximum likelihood estimator. See Hansen et al. (2004) and Heckman et al. (2006) for more details.

I use MCMC techniques to obtain draws from the posterior distribution. Starting with a vector of initial parameters drawn from the transition kernel, I use Gibbs Sampling as the algorithm to create a Markov Chain such that as thesize of the sequence increases ($n \to \infty$) the limiting distribution is the posterior. After convergence is achieved and a burning period of 30,000, I make 1,000 draws from the posterior distribution of the parameters to compute the mean and standard errors of the parameters of interest. More details can be found in the Appendix.

3.6 Results

This section presents four main results. First, I show that unobserved abilities are different from the observed test scores. In particular, the results from a variance decomposition demonstrate that unobserved abilities explain a large fraction of the variance but a significant fraction remains unexplained. In addition, the distributions of test scores and abilities differ significantly for all three dimensions of ability. Moreover, for mechanical ability, the implied sorting into schooling and occupations is completely different when using just the observed test scores.

Second, all three abilities affect schooling, occupational choices and wages. Comparing the magnitude of the effects, cognitive ability has the largest effect on schooling, occupational choice, and average returns.[21] Third, I find a great deal of variation in the size of the economic returns to each dimension of ability. Cognitive ability is highly rewarded in high schooling-abstract occupations while the largest returns to socio-emotional ability are found in low-schooling abstract and high schooling-manual occupations.

Finally, the largest returns to mechanical ability are found in manual occupations, both at high and low levels of schooling.

My results confirm the unique nature of mechanical ability. Unlike standard constructs, it reduces the probability of seeking education beyond high school and the probability of choosing abstract occupations, both of which are associated with higher pay. At the same time, it is positively rewarded in the labor market. This

[21] As I will show in the last section, this is explained by the large effect that cognitive ability has on schooling decisions.

is explained by the large economic returns within manual occupations. Goodness of fit test are passed and the three factors are needed in order to fit the data on wages[22].

3.6.1 Unobserved Abilities

This paper treats observed cognitive, socio-emotional, and mechanical test scores as the outcomes of a process that has as inputs family background, schooling at the time of the test and unobserved abilities. Table 3.2 presents the coefficients on unobserved abilities for each of the tests used. For identification purposes, one loading for each unobserved ability is set to one. The remaining loadings are interpreted in relation to the loading set as the numeraire (for details see Carneiro et al., 2003). The selected numeraires are arithmetic reasoning, mechanical comprehension and the Rosenberg self-esteem scale for cognitive, mechanical and socio-emotional abilities respectively.

Test Scores Variance Decomposition

To analyze the relative importance of each dimension of unobserved ability in explaining test scores, Figure 3.4 presents the variance decomposition of the measurement system. The results show the contribution of observed variables, latent abilities, and error terms as determinants of the variance of each test score.

The variance decomposition illustrates the large size of the unexplained component and highlights the consequences of using observed test scores as proxies for

[22]See Appendix 2 for the estimates and dicussion.

unobserved abilities. The contribution of observed variables to the variance of the test scores is never more than 24 percent. After controlling for the latent variables, the error term is still large but I am able to explain a much higher percentage of the total variance, between 52 and 84 percent. The one exception is the Rotter Scale, where I am only able to explain 14 percent of the variance.

For the three mechanical tests (Auto, Mech. C, and Electronics), both cognitive and mechanical abilities influence the scores. While cognitive ability has lower loadings compared to mechanical ability (see Table 3.2), the variance decomposition shows that both abilities are important determinants of the variance in the observed scores.

In particular, for mechanical comprehension, cognitive explains 19 percent of the variance while mechanical explains 27 percent. For auto shop information and electronics information, cognitive explains 18 and 19 percent respectively while mechanical ability explains 52 and 32 percent of the test score variance. (Disaggregation not shown in the Figure).

Distribution of Abilities and Sorting

As discussed in the previous section, observed test scores and unobserved abilities are different. In this section I use the estimated parameters for the distribution of each ability to estimate the distribution of cognitive, socio-emotional, and mechanical abilities.[23] I show that the distribution of abilities is very different to the distribution of test scores. For mechanical ability, accounting for this difference is

[23] The estimated parameters are presented in Table 3.21 in the Appendix

especially important as the implied sorting into schooling and occupation is completely different when using only observed test scores. The standard deviation and covariance of the simulated distribution for each ability are displayed in Table 4.6.

Figure 3.5 presents the comparison of the cumulative distribution of test scores and abilities (factors) for cognitive and socio-emotional ability. For socio-emotional ability the estimated distribution is bimodal, a characteristic that is not observed when using the test score. Although the distributions are different, the sorting into schooling and occupations is similar. In particular, for both observed test scores and unobserved abilities, the cumulative distribution function (cdf) for people with high education stochastically dominates the cdf curve for people with low schooling. Similarly, the cdf for people in abstract occupations stochastically dominates the cumulative distribution function for those in manual occupations. As a consequence, people with higher levels of ability tend to sort into high levels of education and abstract occupations.

However, for mechanical ability the relationship is reversed. The distribution of the estimated factor implies that people with high levels of mechanical ability choose low education and manual occupations. The cdf of the estimated ability for people in abstract occupations is stochastically dominated by the cdf curve for those in manual occupations (see Figure 3.5). As a consequence, for mechanical ability, the sorting implied by the estimated factor and the observed test scores is completely different, both in terms of schooling (not shown in the Figure) and occupation.

3.6.2 Effect of Abilities on Schooling and Occupational Decisions

Given the nonlinear and multidimensional nature of the model, the best way to understand the results is through simulation. This section presents the simulated effect of increasing each dimension ability by one standard deviation on schooling, occupation, and wages.

Schooling

The decision of continuing education beyond high school is mainly influenced by cognitive ability. Mechanical ability negatively impacts this likelihood and socio-emotional ability has a positive but small effect.

More specifically, a one standard deviation increase in cognitive ability increases the probability of having a higher education by 25 percentage points while for socio-emotional ability, it leads to a 0.3 percentage point increase in that probability. For mechanical ability, a one standard deviation increase reduces the probability of having a higher education in 11 percentage points.[24]

Occupation

As discussed in Section 3.4, the model allows for the occupational decision choice set to vary with on prior schooling choices. Table 3.5 presents the effect of increasing each dimension of ability by one standard deviation on the probability of working in an abstract occupation. The first two rows show the unconditional probabilities

[24] As presented in Table 3.1 the average probability of high education in my NLSY79 sample is 0.31

(Ds_0 and Ds_1); the third row presents the effect on the probability of abstract conditional on schooling decisions.

Cognitive ability has a positive and sizable effect on the probability of selecting an abstract occupation. As expected, the effect on the probability of selecting an abstract occupation is increasing in educational attainment. The total effect is even higher because of the large effect of cognitive ability on schooling. A one standard deviation increase in cognitive ability increases the probability of selecting an abstract occupation by 20.2 percentage points. Socio-emotional ability has a positive impact on the probability of being in an abstract occupation for low schooling levels but a small effect for high schooling levels. The total effect is a 3.8 percentage point increase in the probability of working in an abstract occupation.

The effect of mechanical ability is negative in all cases, but the magnitude is smaller for low schooling. The total net effect is a 8.2 percentage point reduction in the probability of working in an abstract occupation with a one standard deviation increase in mechanical ability.

Wages

In this section I show that the differences in returns to all three abilities across occupations are sizable. The average returns mask these differences, especially in the case of mechanical ability. Mechanical ability is particularly interesting because it is associated with the choices that lead to the lowest wages in the sample: low schooling (high school degree or less) and the choice of manual occupations (see estimated hourly wages in Table 3.6).

As expected, manual occupations offer the highest returns to mechanical ability both at low and high schooling levels. It is particularly interesting to contrast the high returns to mechanical ability with the negative returns to cognitive ability in high schooling-manual occupations. Cognitive ability is rewarded the most in high schooling-abstract occupations where the returns to mechanical are negative and the returns to socio-emotional are negligible.

The abscence of a positive wage response to cognitive ability in manual jobs given the choice of high schooling is surprising because it indicates that more cognitive ability reduces productivity at these type of jobs.[25]

First, it is important to note that most of the studies that compute the returns to ability do not analyze them by schooling and occupation, so there are no other studies to compare this result. However, few studies have found negative returns to measures of cognitive ability for young workers in high levels of scholing (Bishop, 1991; Hause, 1972). Second, negative returns may be the result of a perverse interaction between the requirements of these jobs and the methods prefered by people with high cognitive (arithmentic and verbal) skills to solve problems and follow instructions. People with high arithmentic skills may be more inclined to perform calculations and solve equations in situations where the most efficient strategy is to follow instructions or get the whole picture of the functioning of the instrument or machine they work with. (See Carpenter and Just,1989 for a description of the

[25]From the list of jobs that I classify as manual according to the information from the ONET, the ones that require high schooling levels are for example: electrical and mechanical engineers, technicians (Biological, chemical, and alike), Air traffic controllers, among others

most efficient strategies to solve mechanical problems). A more detailed model is required to fully explain the reasons behind the negative sign.

Finally, the highest returns to socio-emotional ability are for low schooling-manual occupations, followed by high schooling manual occupations. Table 3.7 presents the returns as the effect on log wages of a one standard deviation increase on each dimension of ability.

3.7 Decomposition of the Effect of Ability

In this section I discuss the results of the previous section by decomposing the effect of abilities. Since the model allows for the occupational decision to vary depending on prior schooling choices, I decompose the observed effect of abilities into the fraction explained by the effect through changes in the schooling decision and the direct effect through changing the probability of choosing abstract occupations given an education level.

Similarly, in the case of wages, I decompose the total effect into the fraction explained by changes in the schooling decision (both by signaling and greater productivity), changes in occupational choices, and the direct on-the-job productivity effect. The latter effect reffers to the direct effect of ability on wages that does not operate through schooling degrees/knowledge or occupational differences.

All three abilities have multiple, independent and heterogeneous effects. They significantly affect all three stages: schooling, occupation and wages. All three abilities have a direct impact after holding fixed the indirect channels. The effect of each dimension of ability is different in magnitude and in composition.

This evidence suggests that the large observed effects of cognitive ability on occupation and wages are mainly the result of the large effect it has on schooling attainment.

3.7.1 Decomposition of the Effect on the Occupation Decision

Any change in the vector of abilities would have two effects: a direct effect on the probability of selecting an abstract occupation and an indirect effect through the change in the probability of attaining high schooling. The results indicate that these effects are heterogeneous. A large fraction of the effect of cognitive ability is explained by the effect on schooling while for mechanical and socio-emotional the direct effect is more important.

Letting, ΔDs be the total effect on the probability of choosing an abstract occupation after a one standard deviation increase in one of the three abilities. I simulate the effect of each ability separately so $\theta' = \theta + sd(\theta)$ refers to one ability and assumes the other two abilities are in their original levels.

$$\Delta Ds = \bar{D}s(\theta') - \bar{D}s(\theta)$$

For each individual the observed occupational choice depends on previous schooling decisions

$$Ds_i = D(\theta_i)Ds_1(\theta_i) + (1 - D(\theta_i)) * Ds_0(\theta_i)$$

$$D's_i = D(\theta'_i)Ds_1(\theta'_i) + (1 - D(\theta'_i)) * Ds_0(\theta'_i)$$

In this context the total effect can be decomposed in three parts:

$$\Delta Ds = \Delta D(\bar{D}s_1 - \bar{D}s_0) \tag{3.1}$$

$$+ \bar{D}\Delta Ds_1 + (1 - \bar{D})\Delta Ds_0 \tag{3.2}$$

$$+ \Delta D(\Delta Ds_1 - \Delta Ds_0 \tag{3.3}$$

The effect through schooling (1), assuming that abilities only affect schooling decisions but not occupation decisions conditional on schooling; the effect through occupation (2), which captures how abilities affect occupational choices by changing occupational decisions holding schooling decisions fixed, and the joint effect (3) which accounts for individuals that would select one occupation if education is low and a different occupation if education is high. I refer to this as the joint effect because we only observe an effect if we allow abilities to affect both schooling and occupational choices.

Table 3.8 presents the results from the decomposition of the total effect of abilities on occupational choices. One standard deviation increase in cognitive ability is associated with a 20.2 percentage point increase in the probability of choosing an abstract occupation. A large fraction of this effect is explained by the indirect effect of cognitive ability of schooling choice. In fact, almost 37 percent is explained by cognitive ability increasing the probability of achieving high schooling. Less than half of the effect comes through a direct occupational effect and the remaining 15 percent is explained by the joint effect.

For mechanical ability, the fraction of the effect explained by schooling is larger; almost 45 percent of the 8.2 percentage point decrease in the probability of working in an abstract occupation can be attributed to the reduction in schooling. socio-emotional ability, on the other hand, affects occupation mainly by impacting the abstract/manual occupation decision once scholing is fixed.

3.7.2 Decomposition of the Effect on Wages

In this section, I present the different mechanisms through which skills increase wages. The effect of cognitive ability on wages operates mainly though increasing schooling. socio-emotional and mechanical ability increase wages largely by through their on-the-job productivity enhancement once the occupation and schooling choices have been made. The effect of abilities on wages can be decomposed into four main components: the indirect effect through schooling, the indirect effect through occupation, a direct effect through on-the-job productivity and a joint effect.

Letting ΔW be the total effect on wages after a one standard deviation increase in ability. I simulate the effect of each ability separately so $\theta' = \theta + sd(\theta)$ refers to one ability and assumes the other two abilities are in their original levels.

$$\triangle W = LnW(\theta') - LnW(\theta)$$

For each individual the observed log wage is a function of schooling and occupations:

$$
\begin{aligned}
LnW_i(\theta) &= D(\theta_i) * [Ds_1(\theta_i)w_{11}(\theta_i) + (1 - Ds_1(\theta_i))w_{10}(\theta_i)] + \\
&\quad (1 - D(\theta_i))[Ds_0(\theta_i)w_{01}(\theta_i) + (1 - Ds_0(\theta_i))w_{00}(\theta_i)]
\end{aligned}
$$

The total effect can be decomposed into four parts:

$$
\begin{aligned}
\triangle W &= \triangle D \times (w_1 - w_0) && (3.4) \\
&+ D\left(\triangle D_{s1}\right)\left(w_{11} - w_{10}\right) + (1 - D)\left(\triangle D_{s0}\right)\left(w_{01} - w_{00}\right) && (3.5) \\
&+ D\left\{D_{s1}\left(\frac{\partial \mathbf{w_{11}}}{\partial \theta}\right) + (1 - D_{s1})\left(\frac{\partial \mathbf{w_{10}}}{\partial \theta}\right)\right\} && (3.6) \\
&+ (1 - D)\left\{D_{s0}\left(\frac{\partial \mathbf{w_{01}}}{\partial \theta}\right) + (1 - D_{s0})\left(\frac{\partial \mathbf{w_{00}}}{\partial \theta}\right)\right\} \\
&+ \triangle D \triangle Ds \triangle w_{ij} && (3.7)
\end{aligned}
$$

The effect through schooling (3.4), assuming that abilities only affect schooling decisions but not occupation decisions conditional on schooling; the effect through occupation (3.5), which determines how abilities affect wages by changing occupational decisions but holding schooling decisions fixed; the direct productivity effect (3.6), holding constant the original schooling and occupation decisions; and finally, the joint effect (3.7), which accounts for individuals who would select into different occupation types depending on their education level. I refer to this latter effect as the joint effect because we only observe an effect if we allow abilities to affect all three decision margins (schooling, occupational choices, and wages).

Table 3.9 presents the results from the decomposition. Cognitive ability affects wages mainly by changing schooling, though the there is also a significant productivity effect. In particular, 35 and 33 percent of the observed change in wages resulting from a one standard deviation increase in cognitive ability are explained by the schooling and productivity channels.

In contrast, for mechanical and socio-emotional ability the main channel is the direct, productivity effect. Nearly 87 percent of the observed effect of the socio-emotional factor on wages can be explained by productivity increases. Only 8 and 3 percent are explained by changes in schooling and occupation choices.

For mechanical ability the direct productivity effect is also the strongest. However, mechanical ability is unique in that a one standard deviation increase in mechanical ability has a negative effect on the probability of achieving high schooling and the probability of choosing abstract occupations. Despite this, the direct productivity gains associated with an increase in mechanical ability are so large that they entirely compensate for the negative, indirect effects at the schooling and occupation choice margins.

3.8 Conclusions

In this paper, I analyze the effect of multiple dimensions of pre-labor market abilities on early occupational choices and wages, while taking into account that education decisions are endogenous. My analysis incorporates mechanical ability as an overlooked dimension that, jointly with the other facets of ability explains schooling and occupational decisions as well as labor market outcomes.

In addition, I classify occupations according to their core task requirements which allows to directly associate worker's ability with the activities required at the job. This classification is simple as the standard blue/white-collar classification but it does a better job in classifying some occupations more appropriately. As a result, the proposed classification captures a large fraction of the observed variance in wages.

Finally, by modelling the relationship between schooling, occupations, and wages simultaneously, I identify the main channels through which unobserved initial abilities affect outcomes.

Using the NLSY79, I show that all three dimensions of ability have multiple, heterogeneous, and independent roles. Together, they determine the sorting of workers into schooling and occupations. Cognitive and socio-emotional ability are associated with the choice of high levels of schooling and abstract occupations, while mechanical ability is correlated with the choice of low schooling levels and manual occupations.

A sizable fraction of the effect of pre-labor market abilities on occupational choice is driven by their indirect effects through schooling. Nearly 40 percent of the total effect for cognitive and mechanical ability and 25 percent of the total effect for socio-emotional ability are explained by schooling choices. This indirect effect presumably captures how schooling choices change the choice set of occupations available to workers.

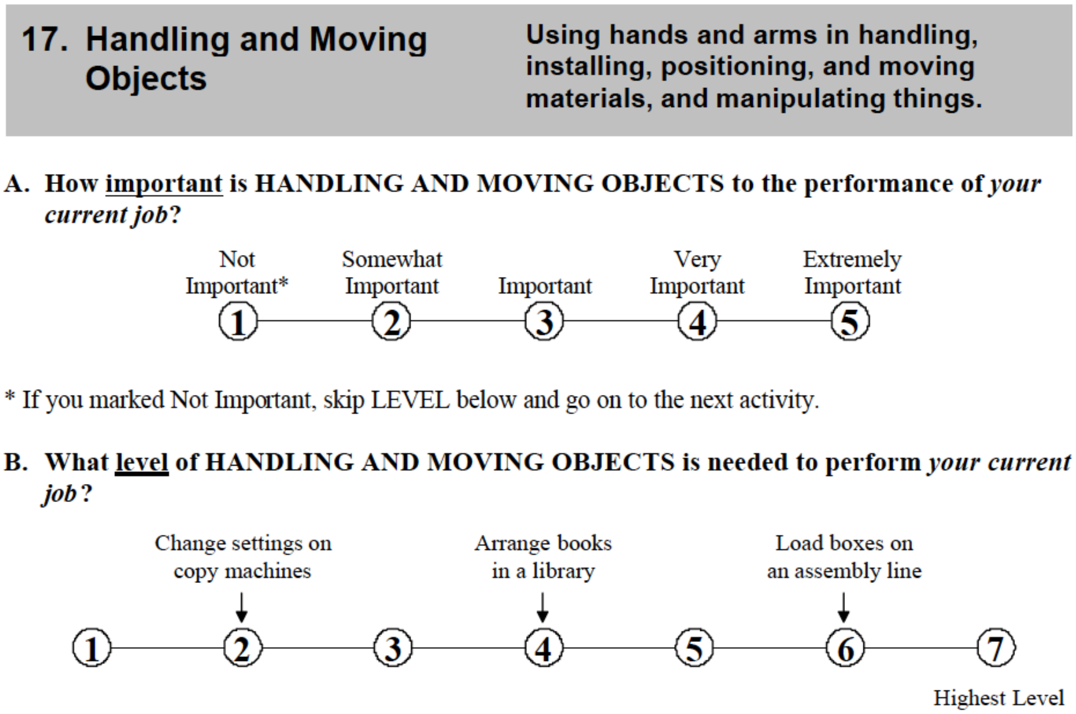
All three skills increase average wages. A one standard deviation increase in cognitive, socio-emotional, and mechanical skills leads to a 12 percent, 6 percent, and 2.7 percent wage increase, respectively.

Moreover, most of the effect of ability of wages remains after discounting the effect through schooling and occupation, what we call the productivity effect. For cognitive skills, the productivity effect represents 33 percent of the total effect, while another 35 percent of the estimated effect is explained by increasing schooling levels. For socio-emotional and mechanical skills, the direct on-the-job productivity effect is the main channel.

Finally, I demonstrate that mechanical ability implies a different profile of schooling and occupational choices and labor market outcomes. Mechanical ability is associated with lower schooling levels and the choice of a manual occupation, but it also has a large, positive effect on wages through its effect on productivity. In fact, the productivity effect of mechanical ability is so large that it completely compensates for the negative, indirect wage effects resulting from the choice of lower schooling levels and manual occupations.

3.9 Tables and Figures

Figure 3.1: Sample Question from O*NET Questionnaire



17. Handling and Moving Objects

Using hands and arms in handling, installing, positioning, and moving materials, and manipulating things.

A. How important is HANDLING AND MOVING OBJECTS to the performance of *your current job*?

Not Important*	Somewhat Important	Important	Very Important	Extremely Important
1	2	3	4	5

* If you marked Not Important, skip LEVEL below and go on to the next activity.

B. What level of HANDLING AND MOVING OBJECTS is needed to perform *your current job*?

	Change settings on copy machines		Arrange books in a library		Load boxes on an assembly line	Highest Level
1	2	3	4	5	6	7

Figure 3.2: Manual and Abstract Composite Task Measures for Blue and White-collar Occupations

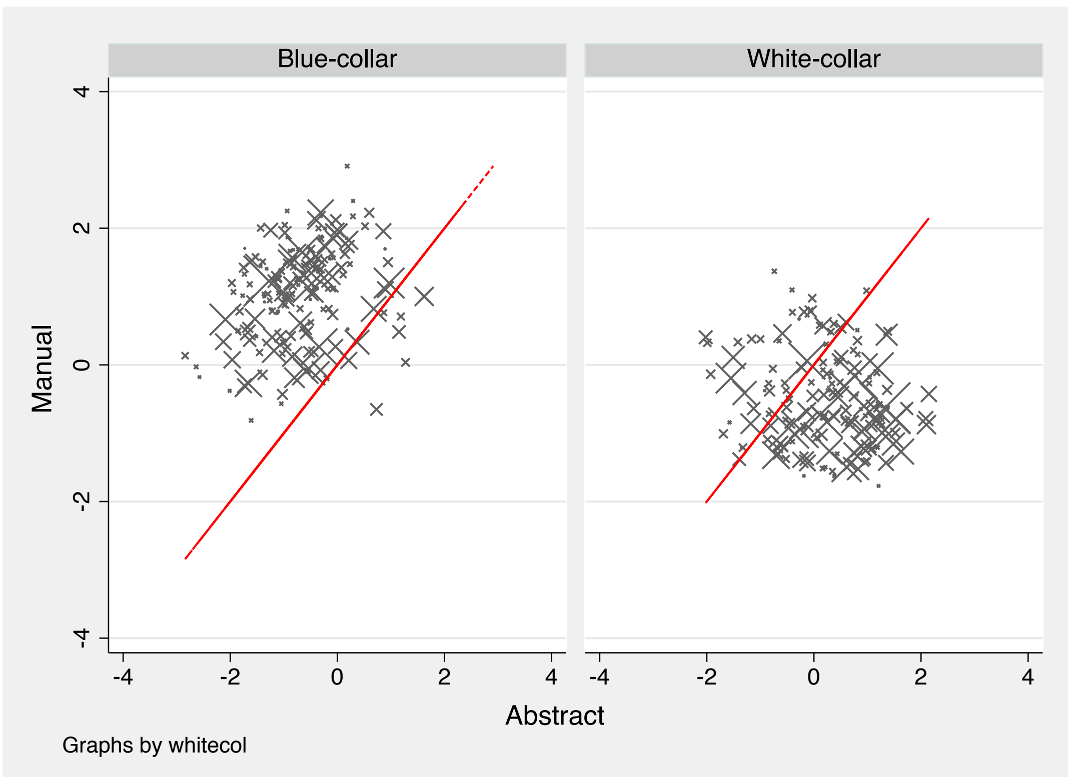
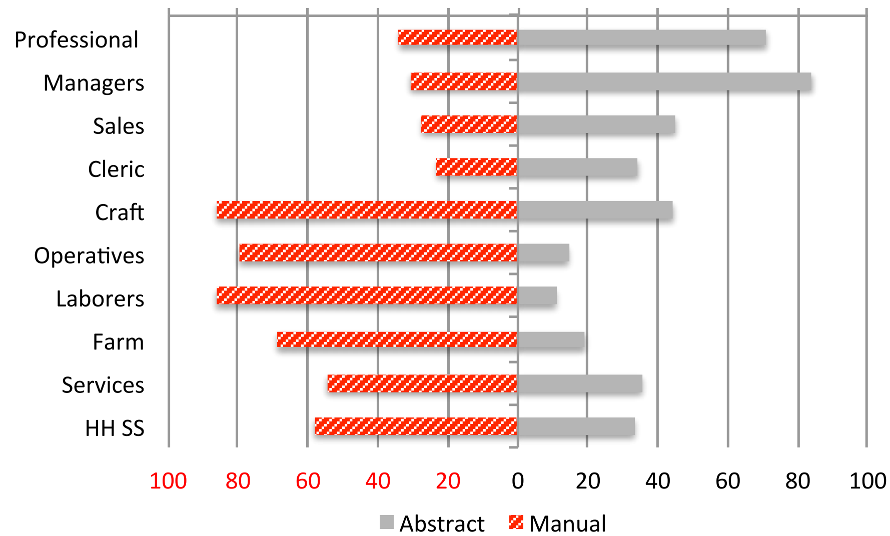
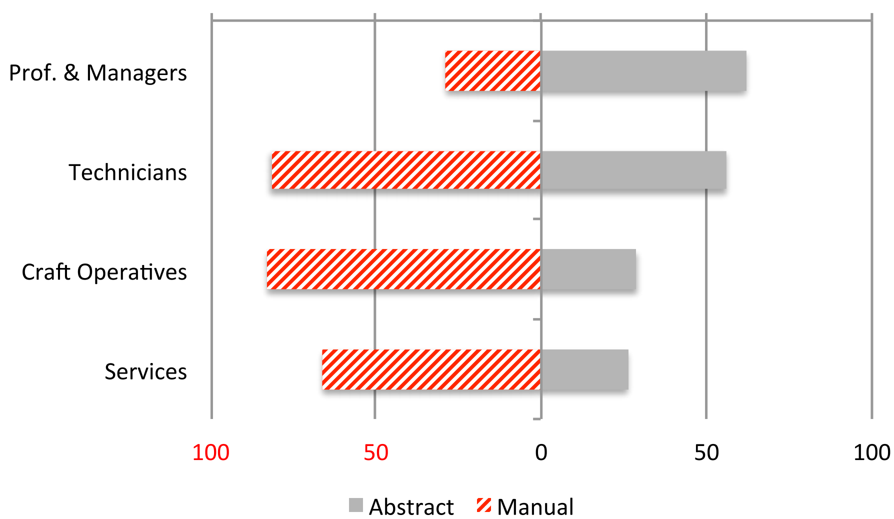


Figure 3.3: Comparison Standard Classification -Manual/Abstract Classification



(a) 10 categories



(b) 4 categories

Note: The figure presents the centiles of the distribution of abstract and manual measures associated to a group of occupations classified in 10 and 4 categories respectively.

Table 3.1: Descriptive Statistics

Variable	Mean	(Std. Dev.)	Min.	Max.	N
LogHourly wage age 26	2.734	(0.468)	0.963	4.251	1402
More than high school	0.315	(0.465)	0	1	1655
White-collar occupation	0.353	(0.478)	0	1	1449
Abstract occupation	0.3	(0.459)	0	1	1435
AFQT	0.009	(0.997)	-2.868	2.011	1655
Mechanical	0.011	(0.996)	-3.008	1.989	1655
NonCognitive	0.007	(1)	-3.031	2.499	1655
Northeast residence	0.167	(0.373)	0	1	1655
Northcentral residence	0.317	(0.466)	0	1	1655
South residence	0.276	(0.447)	0	1	1655
West residence	0.156	(0.363)	0	1	1655
1983-86	0.238	(0.426)	0	1	1655
1987-89	0.434	(0.496)	0	1	1655
1990-1993	0.327	(0.469)	0	1	1655
Family Income in 1979 (thousands)	21.045	(11.748)	0	75.001	1655
Number of siblings 1979	3.04	(1.966)	0	13	1655
Mother's highest grade completed	11.152	(3.347)	0	20	1655
Father's highest grade completed	11.165	(4.142)	0	20	1655
Living in urban area at age 14	0.723	(0.448)	0	1	1655
Living in the south at age 14	0.267	(0.443)	0	1	1655
Education at the time of the test	10.998	(1.195)	6	11	1655

Figure 3.4: Variance Decomposition Test Scores

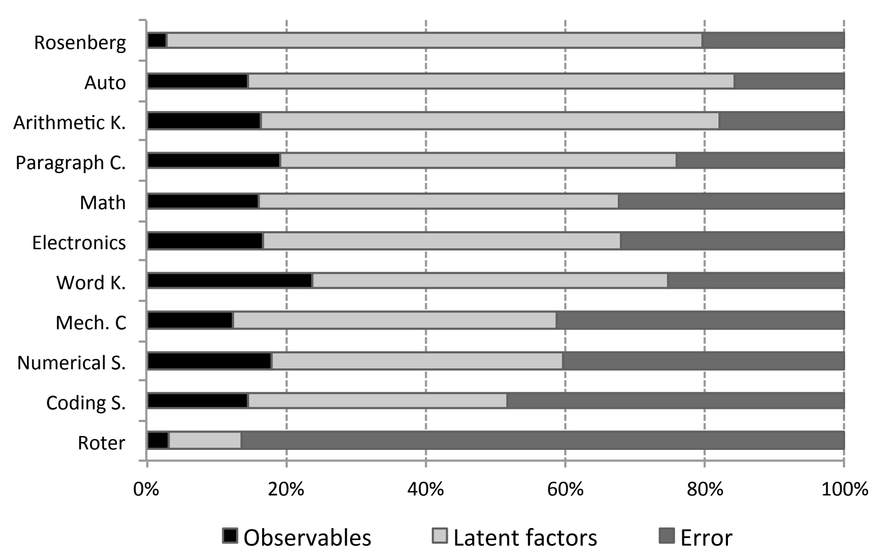


Table 3.2: Loadings of Abilities in Test Scores

	Cognitive	**Mechanical**	**Socio-emotional**
Auto	0	1	
SE	0.03	0.07	
Electronics	0.45	0.61	
SE	0.03	0.05	
Mech. C	0.44	0.64	
SE	0.03		
Arithmetic K.	1.00		
SE			
Math	0.95		
SE	0.00		
Word K.	0.91		
SE	0.03		
Paragraph C.	0.94		
SE	0.03		
Numerical S.	0.77		
SE	0.03		
Coding S.	0.70		
SE	0.03		
Rotter			0.26
SE			0.03
Rosenberg			1.00
SE			

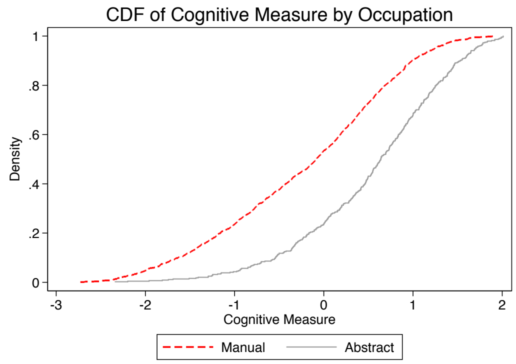
Note: This table presents estimates from the model. Since the model is estimated using Bayesian Methods, they represent the mean estimates over 1,000 iterations after discarting the first 30,000. The computation of standard errors is explained in appendix B. All regressions include family background controls (mother's and father's education, number of siblings, a dummy for broken family at age 14, family income in 1979), schooling level at the time of the test, year dummies and geographical controls for region and urban residence at the age of 14.

Table 3.3: Simulated Parameters of the Distribution of Ability

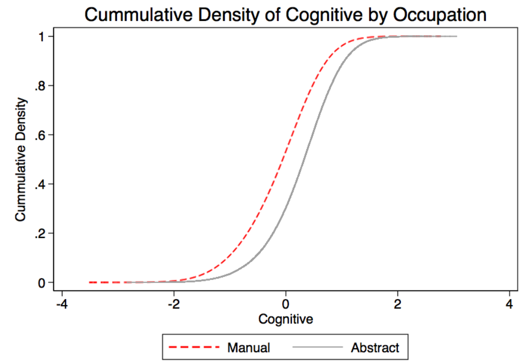
	Simulated
$SD(\theta^c)$	0.73***
$SD(\theta^m)$	0.81***
$SD(\theta^s)$	0.87***
$Cov(\theta^c, \theta^m)$	0.31***
ρ_{θ^c,θ^m}	0.55 * **

Note: Results simulated from the estimates of the model and our NLSY79 sample.

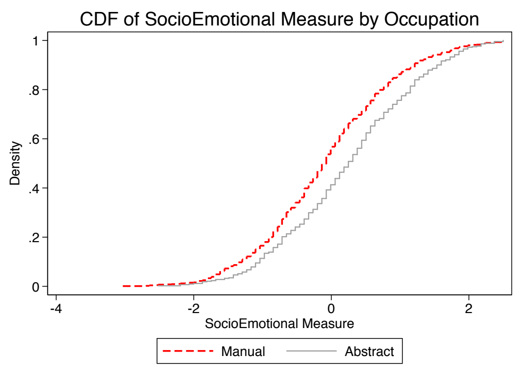
Figure 3.5: Distribution of Test Scores and Abilities by Occupation: Cognitive and Socio-emotional



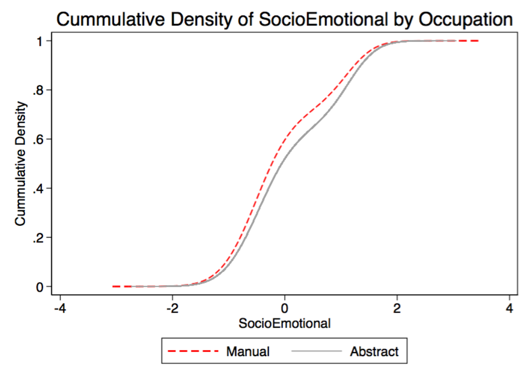
(a) Test



(b) Ability



(a) Test



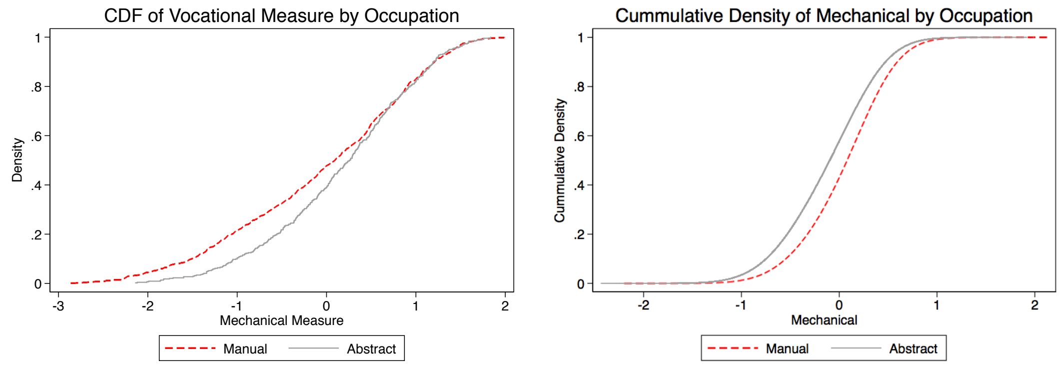
(b) Ability

Note: The cognitive measure (test score) is an average of standarized scores for arithmetic reasoning, word knowledge, paragraph comprehension, mathematics knowledge, numerical operations and coding speed sections of the ASVAB. Socio-emotional test score is an average of the scores in two tests: Rotter Locus of Control Scale and Rosenberg Self-Esteem Scale. The distribution of the factors (unobserved abilities) comes from a simulation using the estimated parameters from the model.

Table 3.4: Simulated Effect of Abilities on Schooling Decisions

ΔD	Pr(more than HS)
Cognitive	0.25 (0.013)***
Mechanical	-0.11 (0.011)***
Socio-emotional	0.03 (0.006)***

Figure 3.6: Distribution of Test Scores and Abilities by Occupation: Mechanical Ability



(a) Test (b) Ability

Note: The measure for mechanical ability (test score) is an average of standarized scores for auto and shop information, mechanical comprehension and electronics information sections of the ASVAB. The distribution of the factor (unobserved abilitiy) comes from a simulation using the estimated parameters from the model.

Table 3.5: Estimated Marginal Effects: Probability of Abstract Occupation

Change in pp	**Cognitive**	**Mechanical**	**Socio-emotional**
Ds_0	6.5%	-2.4%	3.9%
(SE)	(0.007)***	(0.005)***	(0.006)***
Ds_1	18.2%	-12.2%	1.2%
(SE)	(0.011)***	(0.010)***	(0.004)
Abstract Ds	20.2%	-8.2%	3.8%
(SE)	(0.015)***	(0.009)***	(0.006)***

Note: The probability of having an abstract occupation is 0.3, 0.151 for people with high school completed or less and 0.616 for individuals with education beyond high school.

Table 3.6: Estimated Log Hourly Wages by Schooling and Occupation

Schooling	**Manual**	**Abstract**	**Total**
Low	15.8	17.8	16.1
High	18.5	20.9	20.0
Total	16.3	19.8	17.3

Table 3.7: Estimated Marginal Effects on Log Wages by Occupation Given Schooling

Log (wage)	Cognitive	Mechanical	Socio-emotional
W Manual-Low	3.8%	6.7%	4.9%
SE	(0.001)***	(0.001)***	(0.001)***
W Abstract-Low	2.3%	3.6%	11.3%
SE	(0.003)***	(0.002)***	(0.002)***
W Manual-High	-10.4%	6.0%	8.5%
se	(0.003)***	(0.003)***	(0.002)***
W Abstract-High	16.4%	-4.8%	1.2%
SE	(0.002)***	(0.002)***	(0.012)
Total W	12.1%	2.7%	6.2%
SE	(0.014)***	(0.009)***	(0.006)***

Table 3.8: Decomposition of the Effect of Abilities on Occupation

Change in pp	Cognitive	Mechanical	Socio-emotional
(1) Effect through Schooling	7.4	-3.7	0.9
	(0.012)***	(0.008)***	(0.005)*
(2) Effect through Occupation	9.7	-5.7	2.9
	(0.010)***	(0.008)***	(0.006)***
(3) Joint Effect	3.1	1.2	-0.1
(1)+(2)+(3) Total effect	20.2	-8.2	3.8
	(0.015)***	(0.011)***	(0.008)***

Table 3.9: Decomposition of the Effect of Abilities on Wages

Wages	Cognitive	Mechanical	Socio-emotional
(4) Effect through Schooling	4.2%	-1.8%	0.5%
(se)	(0.011)***	(0.007)***	(0.002)**
(5) Effect through Occupation	1.1%	-0.9%	0.22%
(se)	(0.007)*	(0.006)	(0.001)**
(6) Direct Productivity Effect	4.0%	4.2%	5.4%
(se)	(0.003)***	(0.005)*	(0.002)***
(7) Joint Effect	2.7%	1.2%	0.1%
(se)	(0.003)***	(0.005)*	(0.001)*
(4)+(5)+(6)+(7) Total effect	12.1%	2.7%	6.2%
(se)	(0.003)***	(0.002)***	(0.001)***

3.10 Appendixes

3.10.1 Appendix 1: O*NET

The Occupational Information Network (O*NET) is the successor of the U.S. Department of Labor's Dictionary of Occupational Titles (DOT). The DOT has been criticized for not being representative of all occupations, for not following a standard survey design and for poor data quality. [26] The database identifies, defines, describes and classifies over 950 occupations. The O*NET database is continually updated by surveying a broad range of workers from each occupation. The information that populates the O*NET database is collected from three primary sources: incumbents, occupational experts, and occupational analysts. Targeted job incumbents provide ratings on occupational tasks, generalized work activities (GWA), knowledge, education and training, work styles, and work context areas. Importance and level information regarding the abilities and skills associated with these occupations is collected from occupational analysts.

The information available is organized into six major domains. These are: Worker Characteristics, Worker Requirements, Experience Requirements, Occupation Requirements, Occupational Characteristics, and Occupation-Specific Information. I use information from the Occupational Requirements domain, in particular from two sections: general work activities and work context. The other section is organizational context but it does not contain any information relevant for the present

[26] See Miller (1980) for a critical review of the DOT by the National Research Council http://www.nap.edu/openbook.php?record_id=92&page=217.

Table 3.10: Occupational Requirements: A. Generalized Work Activities

A. Generalized Work Activities	1. Information Input		
	2. Mental Processes	a. Information and Data Processing	Judging the Qualities of Things, Services, or People
			Processing Information
			Evaluating Information to Determine Compliance with Standards
			Analyzing Data or Information
		b. Reasoning and Decision Making	Making Decisions and Solving Problems
			Thinking Creatively
			Updating and Using Relevant Knowledge
			Developing Objectives and Strategies
			Scheduling Work and Activities
			Organizing, Planning, and Prioritizing Work
	3. Work Output	a. Performing Physical and Manual Work Activities	Performing General Physical Activities
			Handling and Moving Objects
			Controlling Machines and Processes
			Operating Vehicles, Mechanized Devices, or Equipment
		b. Performing Complex and Technical Activities	Interacting With Computers
			Drafting, Laying Out, and Specifying Technical Devices, Parts, and Equipment
			Repairing and Maintaining Mechanical Equipment
			Repairing and Maintaining Electronic Equipment
			Documenting/Recording Information
	4. Interacting With Others	a. Communicating and Interacting	Interpreting the Meaning of Information for Others
			Communicating with Supervisors, Peers, or Subordinates
			Communicating with Persons Outside Organization
			Establishing and Maintaining Interpersonal Relationships
			Assisting and Caring for Others
			Selling or Influencing Others
			Resolving Conflicts and Negotiating with Others
			Performing for or Working Directly with the Public
		b. Coordinating, Developing, Managing, and Advising	Coordinating the Work and Activities of Others
			Developing and Building Teams
			Training and Teaching Others
			Guiding, Directing, and Motivating Subordinates
			Coaching and Developing Others
			Provide Consultation and Advice to Others
		c. Administering	Performing Administrative Activities
B. Organizational Context			
C. Work Context	1. Interpersonal Relationships	a. Communication	By Communication Method: Public Speaking, phone, mail, letters, face-to-face
			Contact With Others
		b. Role Relationships	Job Interactions
			Work With Work Group or Team
			Deal With External Customers
			Coordinate or Lead Others
		c. Responsibility for Others	Responsible for Others' Health and Safety
			Responsibility for Outcomes and Results
		d. Conflictual Contact	Frequency of Conflict Situations
			Deal With Unpleasant or Angry People
			Deal With Physically Aggressive People
	2.Physical Work Conditions		
	3. Structural Job Characteristics	a. Criticality of Position	Consequence of Error
			Impact of Decisions
			Freedom to Make Decisions
		b. Routine versus Challenging Work	Degree of Automation
			Importance of Being Exact or Accurate
			Importance of Repeating Same Tasks
			Structured versus Unstructured Work
		c. Competition	Level of Competition
		d. Pace and Scheduling	Time Pressure
			Pace Determined by Speed of Equipment
			Work Schedules
			Duration of Typical Work Week

analysis. The general work activities section contains questions in four topics

I am using the Version 17.0 released on July 2012.

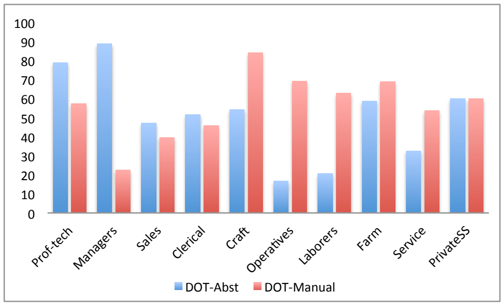
DOT vs. O*NET

Given that I am using data from the NLSY79 it is natural to expect the use of the DOT instead of the O*NET to accurately classify the jobs available for the NLSY79 cohort when they were 26 years old (1984-1991). Actually, the fourth revision of the DOT is dated on 1991 and solves most of the complains about the previous version of the DOT.

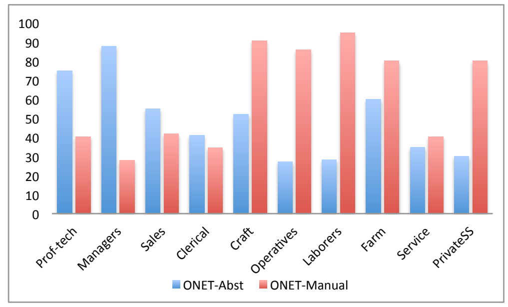
However, I prefer the information from the O*NET over the DOT-4th version for two main reasons: first, the DOT is not representative for services and it was concentrated on manufacturing jobs. In that sense, the O*NET is representative for a larger number of occupations and activities. Second, the DOT, even in its last version, relies too heavily on on-site observations of the jobs by an external individual which reduces the accuracy of the tasks associated with jobs, compared with the alternative of using information from the workers or supervisors.

O*NET is the newer version of the DOT. Although its many advantages the big gap in dates between 2012 and the year when NLSY79 respondents are in their 26 may be problematic because of the compositional changes of jobs and tasks that have taken place during this time. In figure 3.7 I compare the task intensities (in a scale of 1-100) for abstract and manual using both the DOT and the O*NET. Although the intensities have changed, the changes are not large enough to change the classification of jobs in comparable occupations such as manufacture, professionals and managers. In constrast, in service occupations we observe large differences in the classification assoctiated to O*NET and DOT, showing the advantage of the former over the latter.

Figure 3.7: Comparison DOT vs ONET



(a) DOT70



(b) ONET

3.10.2 Appendix 2: Goodness of Fit and Comparison with a Two-Factor Model

This appendix presents evidence on the goodness of fit for schooling decision, occupational choice and hourly wages. It shows that the proposed three-factor model does a better job predicting log wages than a two-factor model that does not include the mechanical factor. Both models predict well schooling and occupational decisions.

Choices

Table 3.11 presents the comparison between the observed choices of schooling and occupation in the data and the resulted averages from the simulation. In both cases a formal Chi-squared goodness of fit test on discrete outcomes evidences the good fit of the model.

Table 3.11: Schooling and Occupational Choices: Observed vs. Simulated

	Observed	Simulated 3f	χ^2	p-value
High Schooling D=1	0.32	0.32	0.08	0.78
Abstract Ds=1	0.30	0.29	0.76	0.38

To compare the performance of the proposed three factor model (cognitive, socio-emotional and mechanical ability) with an alternative two factor model (the standard model with cognitive and socio-emotional ability), I compare them in terms of their goodness of fit. Tables 3.12 and 3.13 present the results of the test for schooling and occupational choices, respectively. The tests cannot reject the null hypothesis which implies that the two models present a good fit with the data. The main difference between the two models is in terms of log wages as I present in the next subsection.

Table 3.12: Goodness of fit 3 factor-model vs 2 factor-model: Schooling

	3 factors	**2 factors**
χ^2	0.08	0.00
p-value	0.78	1.00
Critical at 90%	2.71	2.71
Critical at 95%	3.84	3.84
Critical at 99%	6.63	6.63

Note: The table presents a Chi-squared test for discrete outcomes (Ho:Model=Data.

Table 3.13: Goodness of fit 3 factor-model vs 2 factor-model: Occupation

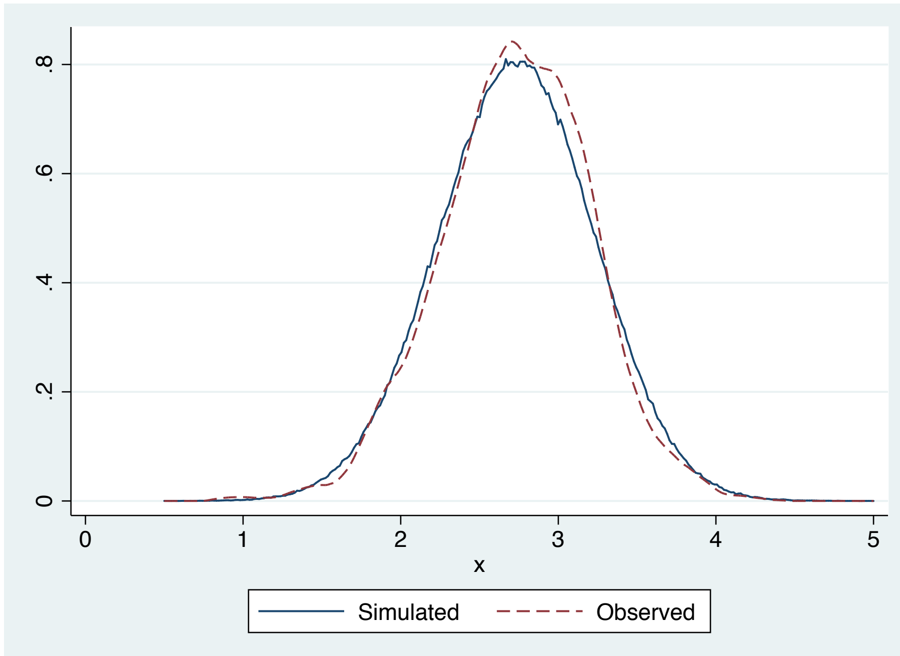
	3 factors	**2 factors**
χ^2	0.76	1.38
p-value	0.38	0.24
Critical at 90%	2.71	2.71
Critical at 95%	3.84	3.84
Critical at 99%	6.63	6.63

Note: The table presents a Chi-squared test for discrete outcomes (Ho:Model=Data.

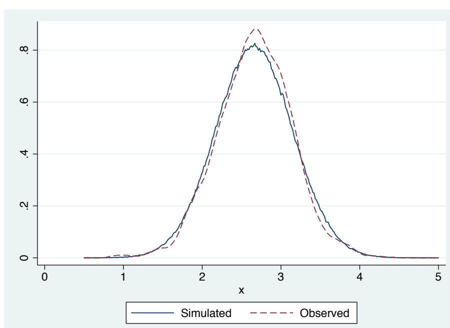
Log wages

Figure 3.8 compares the actual distribution of log wages with the distribution of the simulated log wages for the whole sample (panel a), by schooling level (panels b and c), and by occupational choice (panels d and e). The two distributions are very similar but some differences are evident for individuals in the highest level of education and also for individuals working in abstract occupations. Table 3.14 presents the mean and standard deviation of log wages by schooling, occupational choice and on average. The visual differences observed in the distribution of log wages for individuals with high schooling and individuals working in abstract occupations do not translate into differences in the mean of the log wage but they do translate into larger standard deviations.

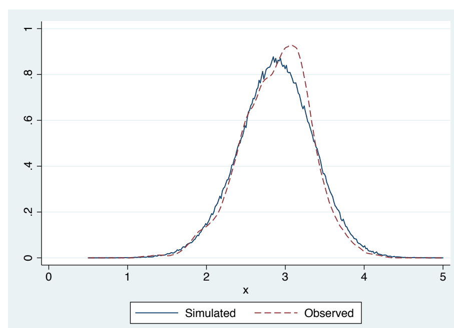
Figure 3.8: Simulated versus Observed Wages



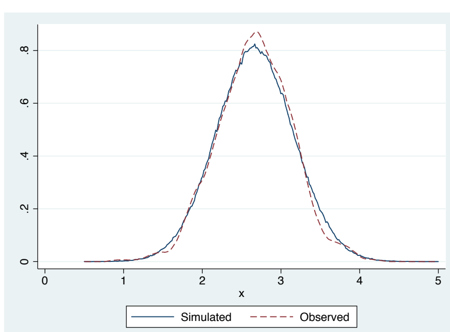
(a) Overall



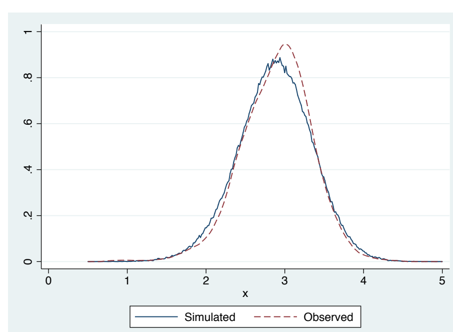
(b) No college



(c) College



(d) Manual



(e) Abstract

Note: The dashed line depicts the actual distribution of log hourly wage in the data while the solid line is computed after simulating a sample of over 1'000.000 individuals using the structure and estimates of the model.

Table 3.14: Log Wages Observed and Simulated by Schooling, Occupational Choice and Average

	Observed	Simulated
Low schooling		
Mean	2.66	2.66
Sd	0.48	0.49
High Schooling		
Mean	2.88	2.88
Sd	0.44	0.47
Manual Occupations		
Mean	2.66	2.67
Sd	0.48	0.49
Abstract Occupations		
Mean	2.90	2.88
Sd	0.43	0.47
Total		
Mean	2.73	2.73
Sd	0.48	0.49

Table 3.15 presents a formal goodness of fit test for log wages wages. The chi-squared test cannot reject the null hypothesis that the simulated distribution of hourly wages is statistically equivalent to the actual distribution observed in the data.

Moreover, the three factor model used is superior than an alternative two factor model that does not take into account mechanical ability. In fact, the two factor model cannot succesfully reproduce the distribution of log hourly wages. Table 3.15 presents the results of the chi-squared goodness of fit test on the simulated distribution of hourly wages that corresponds to a model with three and two factors (only cognitive and socio-emotional). The null hypothesis for the model of two factors is rejected.[27]

Table 3.15: Goodness of Fit: Wage Distribution

	3 factors	**2 factors**
χ^2	51.56	100.93
p-value	0.09	0.00
Critical at 90%	50.66	50.66
Critical at 95%	54.57	54.57

Note: The table presents a Chi-squared test computed using equiprobable bins (Ho:Model=Data).

Tables 3.16 and 3.17 present the same comparison between the 3 factor model used and the alternative 2 factor model for log wages by schooling and occupational choices, respectively. In all cases, the three factor model is superior than the 2 factor model but for manual occupations none of the specifications pass the goodness of fit test.

Table 3.16: Goodness of Fit: Wage Distribution by Schooling

	Low schooling D=0		High schooling D=1	
	3 factors	**2 factors**	**3 factors**	**2 factors**
χ^2	47.99	84.38	33.51	76.08
p-value	0.153	0.00	0.72	0.00
Critical at 90%	50.66			
Critical at 95%	54.57			

Note: The table presents a Chi-squared test computed using equiprobable bins (Ho:Model=Data).

3.10.3 Appendix 3: Estimated Parameters of the Model

[27]It is useful to point out that Heckman et al. (2006) find similar results when computing the

Table 3.17: Goodness of Fit: Wage Distribution by Occupation

	Manual Ds=0		Abstract Ds=1	
	3 factors	**2 factors**	**3 factors**	**2 factors**
x2	66.76	95.05	45.19	76.51
p-value	0.00	0.00	0.23	0.00
Critical at 90%	50.66			
Critical at 95%	54.57			

Note: The table presents a Chi-squared test computed using equiprobable bins (Ho:Model=Data).

Table 3.18: Estimates of the Model: Measurement Equations

	cons	**Sibl**	**Med**	**Fed**	**FamY**	**urban**	**south**	**hgtest**	**coh1**	**coh2**	**coh3**	**c**	**m**	**se**
Auto	-3.09	-0.01	0.00	0.01	0.00	-0.16	-0.23	0.29	0.53	0.34	0.07	0.00	1.00	0.00
SE	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.10	0.09	0.08	0.00	0.00	0.00
Elec.	-3.11	-0.05	0.01	0.02	0.00	-0.06	-0.21	0.27	0.20	0.04	-0.09	0.45	0.61	0.00
SE	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.10	0.09	0.08	0.00	0.00	0.00
Mech.	-2.76	-0.02	0.02	0.02	0.00	-0.12	-0.18	0.23	-0.06	-0.17	-0.18	0.44	0.64	0.00
SE	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.10	0.09	0.08	0.00	0.00	0.00
Ari.	-3.07	-0.01	0.03	0.03	0.01	0.01	-0.19	0.22	-0.30	-0.44	-0.34	1.00	0.00	0.00
SE	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.09	0.09	0.08	0.00	0.00	0.00
Math	-2.34	-0.03	0.03	0.04	0.01	0.02	-0.16	0.14	-0.60	-0.62	-0.25	0.95	0.00	0.00
SE	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.09	0.09	0.08	0.00	0.00	0.00
Word.	-3.49	-0.06	0.03	0.03	0.00	-0.02	-0.17	0.26	-0.10	-0.30	-0.34	0.91	0.00	0.00
SE	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.09	0.09	0.08	0.00	0.00	0.00
Para.	-3.30	-0.04	0.02	0.04	0.00	-0.04	-0.09	0.24	-0.31	-0.39	-0.29	0.94	0.00	0.00
SE	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.10	0.09	0.08	0.00	0.00	0.00
Num.	-3.14	-0.02	0.02	0.03	0.01	0.02	-0.16	0.23	-0.24	-0.41	-0.24	0.77	0.00	0.00
SE	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.10	0.09	0.08	0.00	0.00	0.00
Cod.	-2.90	-0.02	0.02	0.02	0.01	0.01	-0.20	0.22	-0.14	-0.13	-0.19	0.70	0.00	0.00
SE	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.10	0.10	0.08	0.00	0.00	0.00
Rotter	-1.49	0.00	0.00	0.02	0.00	-0.01	-0.04	0.11	0.08	-0.04	-0.08	0.00	0.00	0.26
SE	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.11	0.10	0.08	0.00	0.00	0.00
Rosen.	-1.43	-0.02	0.01	0.01	0.00	0.02	-0.02	0.12	0.18	0.18	0.16	0.00	0.00	1.00
SE	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.10	0.09	0.08	0.00	0.00	0.00

Note: This table presents estimates of the model. Using data from the NLSY79, white males at 26 years old. Since the model is estimated using Bayesian methods, they represent the mean estimates over 1,000 iterations after discarting the first 30,000. The computation of standard errors is explained in appendix B. cons is the constant, Sib is the number of siblings in 1979, Med is the mother's highest grade completed at age 17, Fed is the father's highest grade completed at age 17, FamY is the family income in 1979 in thousands, urban is a dummy variable for living in an urban area at age 14, south is a dummy variable for living in the south at age 14, Coh1 refers to the first cohort (born 57-58), Coh2 refers to the second (born 59-60), Coh3 refers to the last cohort of individuals, those that were born between 61-62, hgtest is the highest grade attended by the time the test was presented and c, m, se refers to the cognitive, mechanical and socio-emotional factors respectively. For space concerns mechanical and socioemotional loadings are collapsed in the same column since they never appear at the same time in any of the test scores specifications. The first three rows refer to the scores in the technical composites of the ASVAB, the next six scores are the tests used to capture cognitive ability and the last two rows are the socio-emotional test scores.

Table 3.19: Estimates of the Model: Schooling Model

Pr(Beyond High School)	Coefficient	SE
Constant	-1.74	0.01
Number of siblings	-0.13	0.00
Mother's highest grade completed	0.05	0.00
Father's highest grade completed	0.11	0.00
Family Income 1979 (thousands)	0.01	0.00
Living in urban area at age 14	0.05	0.00
Living in the south at age 14	-0.22	0.00
Tuition college at age 17	-0.03	0.00
Cognitive	0.96	0.00
Vocational	-0.62	0.00
Socio-emotional	0.14	0.00

Table 3.20: Estimates of the Model: Wages given Schooling and Occupational Choice

Log Wages	Manual | Low S		Abstract | Low S		Manual | High S		Abstract | High S	
	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
Constant	2.77	0.00	2.66	0.01	2.76	0.00	2.82	0.00
Northeast residence	0.04	0.00	0.04	0.00	0.21	0.00	0.10	0.00
Northcentral residence	-0.11	0.00	-0.02	0.00	-0.03	0.00	-0.05	0.00
South residence	-0.17	0.00	-0.06	0.00	-0.14	0.00	0.00	0.00
Cohort1 (Born 57-58)	-0.02	0.00	0.07	0.00	0.13	0.00	0.09	0.01
Cohort3 (Born 61-62)	-0.03	0.00	-0.02	0.00	0.01	0.00	-0.01	0.00
Local Unemployment rate	-0.32	0.02	1.46	0.06	0.06	0.05	0.59	0.03
Cognitive	0.09	0.00	0.07	0.00	-0.08	0.00	0.19	0.00
Vocational	0.08	0.00	0.02	0.00	0.08	0.00	-0.05	0.00
Socio-emotional	0.05	0.00	0.13	0.00	0.08	0.00	0.00	0.00

Table 3.21: Estimated Parameters of the Distribution of Abilities

	Cognitive		Mechanical		Socio-emotional	
	Estimate	SE	Estimate	SE	Estimate	SE
μ_1	0.35	0.14	-0.36	0.11	-0.48	0.07
μ_2	-0.51	0.30	0.36	0.06	1.10	0.11
τ_1	3.76	0.94	5.45	1.29	3.90	1.01
τ_2	2.58	0.81	12.92	2.86	6.46	2.00
ρ	0.56	0.21	0.51	0.11	0.70	0.05
1-ρ	0.44	0.21	0.49	0.11	0.30	0.05

Note: This table presents estimates from the Model. Since the model is estimated using Bayesian Methods, they represent the mean estimates over 1,000 iterations after discarding the first 30,000. The computation of standard errors is explained in Appendix 3. All regressions include family background controls (mother's and father's education, number of siblings, a dummy for broken family at age 14, family income in 1979), year dummies and geographical controls for region and urban residence at the age of 14.

3.10.4 Appendix 4: Robustness Checks

3.10.4.1 Addressing Selection into Vocational Elective Courses

One potential source of contamination of the test scores used to measure mechanical ability is the fact that some high schools offer vocational elective courses on auto shop, mechanics and electronics. In this case, observed differences in performance in mechanical, auto shop and electronics tests may reflect preferences for certain types of topics, extra preparation or even an anticipatory behavior of students planning to drop out in the future and acquiring skills for jobs in these trades.

The most natural way to control for this potential source of contamination is to restrict the sample to the students that have not yet decided on elective courses by the time the test were presented. Since vocational courses were only available to students after completing 8th grade, restricting the sample to students that have not started 9th grade could, at least in principle, solve the problem. However, in the summer of 1980 only a small fraction of the sample has not started 9th grade. In fact, from the original 1,655 males that we use in the estimations only 211 survive after the restriction. With this very small number of observation is not possible to run the model. Moreover, in order to get a correct sample one must restrict not only the grade at the time of the test but also the age to control for those individuals that have not started 9th grade because they have repeated some years.

Fortunately, the survey provides school transcripts for a subsample of the individuals so it is possible to separate individuals that have credits in one of these

Chi-squared test on the sample of 4-year college graduates.

Table 3.22: Comparison of Simulated Effect of Abilities on Schooling Decisions

ΔD	Pr(more than HS) original	Pr(more than HS) notech
Cognitive	0.25 (0.013)***	0.24 (0.013)***
Mechanical	-0.11 (0.011)***	-0.11 (0.009)***
Socio-emotional	0.03 (0.006)***	0.03 (0.005)***

Table 3.23: Comparison of Simulated Effect of Abilities on Occupational Choice

Change in pp	Pr(Abstract) original	Pr(Abstract) notech
Cognitive	20.2% (0.015)***	19.2% (0.013)***
Mechanical	-8.2% (0.009)***	-7.8% (0.006)***
Socio-emotional	3.8% (0.006)***	4.1% (0.007)***

courses from individuals that did not take any of these courses in high school. Using the school transcripts, I can identify 544 individuals that have not taken any of the elective courses related with the topics of the three tests used to identify mechanical ability.[28]

Inclusion Dummy for Technical/Vocational Course Takers The first check of my results consists in including a dummy variable in the measurement equations to identify students that have taken at least one credit of vocational courses and re-run the model. Results don not change. The effect of abilitiy on schooling and occupational decisions does not change in magnitude (see Tables 3.22 and 3.23).

Some minor changes are observed in terms of the effect of each type of ability on wages. In particular, the returns to mechanical and socioemotional ability are lower than the ones estimated in the original regresions, while the returns to cognitive

[28] It is mimportant to note that these individuals have also special characteristics because they are biased against vocational courses in these topics. For example, they have a higher probability of pursuing more education beyond high school, 46.9 vs 31 on avergae for the whole sample.

Table 3.24: Estimated Marginal Effects on Log Wages by Occupation Given Schooling

Log (wage)	**Cognitive**	**Mechanical**	**Socio-emotional**
W Manual-Low	5.1%	6.5%	3.7%
SE	(0.001)***	(0.001)***	(0.001)***
W Abstract-Low	4.4%	3.6%	11.0%
SE	(0.002)***	(0.002)***	(0.002)***
W Manual-High	-9.7%	6.0%	6.6%
se	(0.002)***	(0.002)***	(0.002)***
W Abstract-High	15.3%	-4.5%	0.2%
SE	(0.001)***	(0.002)***	(0.001)**
Total W	12.4%	2.9%	5.0%
SE	(0.012)***	(0.008)**	(0.006)***

Table 3.25: Simulated Parameters of the Distribution of Ability

	Simulated
$SD(\theta^c)$	0.72***
$SD(\theta^m)$	0.71***
$SD(\theta^s)$	0.87***
$Cov(\theta^c, \theta^m)$	0.29***
ρ_{θ^c,θ^m}	$0.57 * **$

Note: Results simulated from the estimates of the model and the NLSY79 sample used for main results.

ability are practically unchanged. Table 3.24 presents the estimated marginal effects on log wages by occupation given schooling choices. This Table compares with Table 3.7 in the text.

The simulated parameters of the distribution of abilty are also very similar to the original estimates, although the variance of mechanical factor us slightly lower.

Subsamples Another alternative is to restrict the sample to individuals that chose not to take any of the vocational courses in high school. This exercise is interesting in the sense that tests the prediction on the model on a sample that, in principle must have either less interest in vocational/courses and/or less early investments on initial mechanical ability. Unfortunately, the resulting sample is not

Table 3.26: Comparison of Simulated Effect of Abilities on Schooling Decisions

ΔD	Pr(more than HS) original	Pr(more than HS) notech
Cognitive	0.25 (0.013)***	0.26 (0.019)***
Mechanical	-0.11 (0.011)***	-0.15 (0.016)***
Socio-emotional	0.03 (0.006)***	0.04 (0.009)***

Table 3.27: Comparinson of Simulated Effect of Abilities on Occupational Choice

Change in pp	Pr(Abstract) original	Pr(Abstract) notech
Cognitive	20.2% (0.015)***	27.3% (0.021)***
Mechanical	-8.2% (0.009)***	-10.6% (0.017)***
Socio-emotional	3.8% (0.006)***	3.0% (0.009)***

large enough to identify differences in wages by both schooling and occupational choices. Tables 3.26 and 3.27 present the results of a simpler version of the model to confirm that the predictions of the original model hold in terms of schooling choices and occupational choices separately.

3.10.4.2 Task Classification vs Blue/White Classification of Occupations

In this section I compare the standard Blue/White-collar classification of occupation with the Manual/Abstract classification and show the advantages of the latter classification of occupation for the analysis.

A simple comparison of the two classifications in terms of the frequencies and average log wages by category reveals small differences. Table 3.28 compares the share of individuals classified as working in abstract occupations with the share working in white-collar occupations. The differences are small, never more than five

Table 3.28: Distribution of Abstract and White-collar by Schooling

Schooling	White	Abstract	t
Low	0.15	0.21	-3.22 ***
High	0.62	0.67	-1.60 *
Total	0.30	0.35	-3.22 ***

Table 3.29: Log-wages in Abstract and White-collar Occupations by Schooling

Schooling	Manual	Blue	t	Abstract	White	t
Low	2.64	2.65	-0.02	2.78	2.74	0.72
High	2.77	2.77	-0.04	2.95	2.94	0.47
Total	2.66	2.67	-0.06	2.90	2.86	1.19

percentage points, but always statistically significant. The oppposite is true for the log-wages comparison as presented in Table 3.29 there are no statistically significant differences between the categories.

However, the Manual/Abstract classification is superior to the standard classification. On the one hand, because it presents two methodological advantages. First, it provides a simple and systematic way of classifying a large number of occupations according to the main job requirements into just two categories. Second, it classifies more appropriately at least three types of occupations: 1) services that are classified as blue collar but are similar in tasks to white-collar jobs, such as high rank police chief, detectives, etc; 2) Technicians whose characteristics may resemble other white collar workers but spend most of their time working with machines as their blue-collar counterparts; and 3) factory operatives, working in industrial laboratories who are classified as blue-collar but performing activities and tasks similar to technicians. See Table 3.30 for some examples of these occupations and Table

Table 3.30: Examples of Specific Occupations

	Schooling	Manual	Abstract
Blue Collar	**Low**	Drivers, construction laborers, carpenters	Protective services, production supervisors
	High	Automotive mechanics, Cooks	Chief of Police High rank detectives
White Collar	**Low**	Some clerks, Cashiers, Musicians	Farm owners, administrators
	High	Technicians, Airplane pilots, bank tellers	Professional, managers, sales

3.31 for the frequencies in each category.

Table 3.31: Differences between Standard and Proposed Classification of Occupations: Number of Observations and Percentage of Standard Category

	Schooling	Manual	Abstract	Total
Blue Collar	**Low**	765	5	783
		99	1	100
	High	141	12	153
		92	8	100
White Collar	**Low**	62	142	204
		30	70	100
	High	36	272	308
		12	88	100
Total	**Low**	827	147	974
		85	15	100
	High	177	284	461
		.38	.62	100

One the other hand, the advantage of the Manual/Abstract classification goes beyond theorical and methodological issues, it also explains a larger percentage of the observed variance in wages when compared with the alternative blue/white-collar classification.

I examine the relationship between the two classification of occupation and wages by regressing workers log hourly wages on their human capital, demographic characteristics, and different occupational classifications. As a benchmark, column 1 in Table 3.32 presents a standard cross-sectional Mincerian wage regression of log hourly wages on human capital and demographic measures. All variables in this regression have the expected signs and magnitudes. The R-squared of this model is equal to 0.139, comparable to standard cross-sectional models estimated using the NLSY79 on a sample of white males in prime age.

Column 2 includes the dummy variable for abstract occupation, which increases the R-squared to 0.149. Column 3 includes a dummy for high schoolig as defined in this document, which increases the R-squared to 0.151. The rest of the Table replicates columns 2 and 3 using the dummy variable for white collar occupation. This latter specification always explain a lower percentage of the observed variance in wages ranging between 10 and 7 percent less. Finally, Table 3.33 presents a similar set of regressions but using a full 236 occupational dummies to have a sense of the maximum amount of variation that could be explained with occupational information, in this case no more than 26 percent.

Table 3.32: OLS Regressions of Log Hourly Wages on Occupation defined by Task or Blue/White-collar, Demographic Variables and Test Scores

	(1)	(2)	(3)	(4)	(5)
Cognitive	0.0824	0.0746	0.0741	0.0782	0.0783
	(0.0189)***	(0.0193)***	(0.0196)***	(0.0193)***	(0.0193)***
Socioemotional	0.0392	0.0369	0.0372	0.0379	0.0379
	(0.0128)***	(0.0129)***	(0.0129)***	(0.0128)***	(0.0128)***
Mechanical	0.0305	0.0340	0.0372	0.0326	0.0326
	(0.0171)*	(0.0173)**	(0.0174)**	(0.0172)*	(0.0172)*
Abstract occ.		0.0802	0.0741		
		(0.0308)***	(0.0314)**		
White-collar occ.				0.0347	0.0350
				(0.0292)**	(0.0294)**
Demographic vars.	Yes	Yes	Yes	Yes	Yes
Schooling	No	No	Yes	No	Yes
Observations	1441	1425	1425	1439	1439
Adjusted R^2	0.139	0.149	0.151	0.138	0.136

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: cognitive is an average of standarized scores for arithmetic reasoning, word knowledge, paragraph comprehension, mathematics knowledge, numerical operations and coding speed sections of the ASVAB, Socio-emotional is an average of the scores in two tests:Rotter Locus of Control Scale and Rosenberg Self-Esteem Scale. Mechanical is an average of standarized scores for auto and shop information, mechanical comprehension and electronics information sections of the ASVAB. Demographic variables included cohort dummies, geographical controls for region, experience and experience squared.

Table 3.33: OLS Regressions of Log Hourly Wages on 236 Occupational Dummies, Demographic Variables and Test Scores

	(1)	(2)	(3)	(4)
Cognitive		0.0806	0.0529	0.0532
		(0.0189)***	(0.0201)***	(0.0201)***
Socioemotional		0.0310	0.0330	0.0335
		(0.0132)**	(0.0132)**	(0.0132)**
Mechanical		0.0256	0.0253	0.0247
		(0.0179)	(0.0184)	(0.0184)
More than high school				-0.0299
				(0.0419)
Demographic vars.	No	No	Yes	Yes
Schooling	No	No	No	Yes
Observations	1442	1442	1441	1441
Adjusted R^2	0.182	0.225	0.246	0.246

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: cognitive is an average of standarized scores for arithmetic reasoning, word knowledge, paragraph comprehension, mathematics knowledge, numerical operations and coding speed sections of the ASVAB, Socio-emotional is an average of the scores in two tests:Rotter Locus of Control Scale and Rosenberg Self-Esteem Scale. Mechanical is an average of standarized scores for auto and shop information, mechanical comprehension and electronics information sections of the ASVAB. Demographic variables included cohort dummies, geographical controls for region, experience and experience squared.

Chapter 4: **One Size Does Not Fit All: Multiple Dimensions of Ability, College Attendance and Wages**

Note: This chapter of the dissertation is coauthored with Sergio Urzúa.

4.1 Introduction

The importance of cognitive and socio-emotional ability in explaining schooling attainment and labor market outcomes has received considerable attention in the literature. Over the last decades, several studies have found that these abilities affect a number of outcomes. In particular, studies have shown that both types of abilities *positively* affect the acquisition of skills and education as well as market productivity as measured by wages. (See Cawley et al., 2001; O'Neill, 1990; Neal and Johnson, 1996; Herrnstein and Murray, 1994; Bowles et al., 2001; Farkas, 2003; Heckman et al., 2006; Urzua, 2008, among others).

But ability is multidimensional in nature and thus, it is reasonable to expect that other dimensions may also affect schooling decisions and labor market outcomes. In fact, economists have recognized that the multidimensionality of ability must be at the "center stage of the theoretical and empirical research on child development, educational attainment and labor market careers" (Altonji, 2010). Also, recent studies in economics, psychology, and other social sciences have been exploring the different components of socio-emotional ability, generally in the form of personality traits (Borghans et al., 2008; Heckman and Kautz, 2013), but less consideration has been given to the exploration of other facets, especially those that might be related to cognition.

This paper investigates a dimension of ability that has been overlooked by economists when analyzing schooling decisions and labor market outcomes. This dimension is related to motor skills, visual motor integration, and potentially to manual dexterity. We label it "mechanical ability".[1]

To analyze the empirical importance of this ability - jointly with the conventional dimensions -, we implement a Roy model of self-selection into college and counter factual adult wages with unobserved heterogeneity. This framework is similar to the setup analyzed in Carneiro et al. (2003) and Heckman et al. (2006), so we follow their identification strategy. In particular, we augment the Roy model with a set of proxy measures containing multiple test scores (measurement system) from which we identify the distribution of a three-dimensional vector of latent abilities: cognitive, socio-emotional and mechanical.

We contribute to the literature by documenting that mechanical ability matters. We show that it affects schooling decisions and labor market outcomes differently than other measures of ability. In particular, using data from the National Longitudinal Study of Youth of 1979 (NLSY79), we show that, like cognitive and

[1]Other papers have studied the importance of aspects connected to the idea of "mechanical ability", and their association with labor market outcomes (see for example Hartog and Sluis, 2010; Yamaguchi, 2012; Boehm, 2013, among others). However, this literature does not simultaneously analyze multiple abilities, schooling decisions and labor market outcomes.

socio-emotional abilities, mechanical ability has a positive economic return, but in contrast to conventional dimensions, it predicts the choice of low levels of schooling. In particular, it reduces the probability of attending four-year college. In this context, this dimension expands the set of abilities explaining differences in human capital and wages in the population.

To identify this ability, we utilize a set of questions from the Armed Services Vocational Aptitude Battery (ASVAB) that has been historically used by the military to determine qualification for enlistment in the United States armed forces. But despite its popularity, only a subset of these questions has been investigated in the literature: the battery of tests used to calculate the Armed Forces Qualification Test (AFQT) score, which is commonly interpreted as a proxy for cognitive ability. This paper highlights the importance of the technical composites of the ASVAB to capture a different dimension of ability.

Our study provides insight into the schooling choices and earnings of individuals conventionally classified as low-ability, but who might be endowed with a high level of mechanical ability. We present evidence that for them, not going to college implies a higher expected hourly wage compared to the expected hourly wage associated with college attendance. This has important implications for public policies promoting general enrollment in four-year colleges.

The paper has six sections. The second section describes the data used and presents reduced-form estimates of the implied effect of mechanical ability on schooling choices and wages, both unconditional and conditional on conventional observed measures of cognitive and socio-emotional ability. Section three contains the details of our augmented Roy model and the estimation strategy. Section four presents the main results. Section five presents a discussion of the implication of our results. Section six concludes.

4.2 Data and Exploratory Analysis

We now turn to the description of our source of information, a brief discussion of the measure of mechanical ability in comparison with conventional measures of ability, and the reduced-form estimates of the effect of mechanical ability on schooling choices and wages both unconditional and conditional on standard measures of ability. The insights from the descriptive analysis are used in two ways: to document that mechanical ability is correlated with schooling decisions differently than standard measures of ability, and to motivate the use of a model to capture the effect of mechanical ability overcoming the main problems associated with the reduced-form estimates.

4.2.1 Data

The National Longitudinal Survey of Youth (NLSY79) is a panel data set of 12,686 individuals born between 1957 and 1964.[2] This survey is designed to represent the population of youth aged 14 to 21 as of December 31 of 1978, and residing in the United States on January 1, 1979. It consists of both a nationally representative cross-section sample and a set of supplemental samples designed to oversample civil-

[2] Includes 2,439 white males,21 percent of total surveyed individuals and 40 percent of the individuals in the cross-sample.

ian blacks, civilian Hispanics, economically disadvantaged Non-Black/ Non-Hispanic youths, and individuals in the military. Data is collected in an annual basis from 1979 to 1994 and biannually until present day.

We use the cross-section sample of white males between the ages of 25 and 30 who were not attending school at the time of the survey and who were, at most, high school graduates at the time of the tests used to measure ability were collected (Survey of 1979 and the summer 1980). We chose to analyze white males in order to have a benchmark to compare our results with previous studies (Heckman et al., 2006; Neal and Johnson, 1996, etc). In addition, we want to abstract from influences that operate differently on various demographic groups. In consequence, our analysis is specific and cannot be generalized to the whole population.

The age selection responds to the interest of analyzing entry level wages abstracting from the cumulative effects of ability on experience and tenure. By the age of 25, more than 97 percent of the sample has reached their maximum level of education. The five-year window is useful to get a smooth average of the first part of the wage profile of the individuals.

From the original sample of 12,686 individuals, 11,406 are civilian, 6,111 belong to the cross-section sample. Nearly 49 percent of that sample are males (2,438 individuals), 1,999 had less than high school complete by the time the ASVAB test was conducted (Summer 1980), out of them just 1,832 individuals are observed at least once between the ages of 25 and 30 and finally, 1,710 were not attending school by the time the survey was conducted. That sample is further reduced for the analysis according to the variables of interest. We got rid of 244 observations that either are high school dropouts or have no information on schooling. We ended up with a sample of 1,466 individuals. Table 2 presents the description of the variables used.

We analyze one schooling choice, four-year college attendance. The variables used to determine college attendance are maximum degree attained by the age of 25 and type of college enrolled. The labor market outcome analyzed is the log of the average of the hourly wages reported between 25 and 30 years old.

For the cognitive and mechanical measures we rely on the (ASVAB) that was conducted in the summer and fall of 1980.[3] This questions are used to compute the AFQT that is used by the military services for enlistment screening and job assignment within the military. This test was administrated to over 90 percent of the members of the NLSY panel (individuals were between 15 and 23 years old at the time of the test). The test is composed by a battery of 10 questions measuring knowledge and skills in the following areas: arithmetic reasoning, word knowledge, paragraph comprehension, numerical operations, coding speed, mathematics knowledge, general science, auto and shop information, mechanical comprehension, and electronics information. The first 6 are used as measures of cognitive ability while the last 3 are measures of mechanical ability.

For measures of socio-emotional ability we use two tests: the Rotter Locus of Control Scale and the Rosenberg Self-Esteem Scale. The Rotter Locus of Control Scale measures the degree of control individuals feel they possess over their life.[4] In

[3]These questions are used to compute the AFQT that is used by the military services for enlistment screening and job assignment within the military.

[4]These measures have been used in the literature as proxies of socio-emotional ability (Heckman et al., 2006)

1979 the NLSY collected a total of four items selected from the 23-item forced choice questionnaire adapted from the 60-item Rotter Adult I-E scale developed by Rotter (1966). As presented in the NLSY79 documentation: "This scale was designed to measure 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 (that is, chance, fate, luck) controls their lives (external control). The scale is scored in the external direction-the higher the score, the more external the individual".[5]

The Rosenberg Self-Esteem Scale, which is based on 10 questions, measures self-esteem: the degree of approval or disapproval towards oneself (Rosenberg, 1965). The scale is short, widely used, and has accumulated evidence of validity and reliability. It contains 10 statements of self-approval and disapproval with which respondents are asked to strongly agree, agree, disagree, or strongly disagree. The scale has proved highly internally consistent, with reliability coefficients that range from .87 (Menaghan, 1990) to .94 (Strocchia-Rivera, 1988), depending on the nature of the NLSY79 sample selected".[6]

4.2.2 Distributions

The tests are used to create a composite measure for each type of ability. For cognitive ability the measure is constructed using an average of the standardized scores for arithmetic reasoning, mathematical knowledge, paragraph comprehension, word

[5]Extracted from http://www.nlsinfo.org/nlsy79/docs/79html/79text/attitude.html.
[6]Ibid.

knowledge, numerical operations, and coding speed. For socio-emotional ability the measure is created as the sum of the average of Rotter and Rosenberg scores. Finally, mechanical ability measurement is constructed as the average of the standardized scores in mechanical comprehension, electronics information, and auto and shop information.

We are mainly interested in the sorting implied by each measure of ability. Figures 4.3 and 4.4 show the cumulative distribution function (cdf) of each measure by schooling choice. For all three measures of ability, the cdf for people with high education stochastically dominates the cdf curve for people with low schooling. As a consequence, people that score higher in these measures of ability tend to sort into high levels of education.

This result is not surprising but in the next section we show that when we control for all three measures, mechanical ability implies a different and interesting behavior, the one that motivates this paper.

4.2.3 Reduced-form Effect on Schooling Choice

To analyze the effect of the mechanical tests on schooling choices we estimate a probit model for the probability of attending 4-year college. All regressions include a set of family background controls, cohort dummies and dummies for region and urban location.

The unconditional effect of the mechanical test on college attainment is positive as it is the effect of cognitive ability, but the magnitude is smaller. Analyzing the marginal effects evaluated at the mean (MEM) presented in Table 4.3 (columns 1 and 2) both cognitive and mechanical tests show a similar pattern in terms of the positive impact on schooling attainment but the effect of AFQT more than doubles of the effect of the measure of mechanical ability.

This result is expected given the sorting implied by the distribution of each measure of ability (scores in the tests) as presented in figures 4.3 and 4.4.

But controlling for AFQT, the effect of the mechanical test on educational attainment is reversed. In particular, the marginal effects evaluated at the mean (MEM) presented in column 3 show that once cognitive and socio-emotional scores are taken into account, one standard deviation increase on the mechanical test decreases the probability of attending a 4-year college in 6.23 percentage points. While the same increase on the cognitive test increases college attendance by 20.6 percentage points. This effect is large considering that in the sample the probability of attending college is 29 percent and the predicted probability at the mean of the observed variables is 22.6.

4.2.4 Reduced-form Effect on Hourly Wages

Analyzing hourly wages, the return to the score in the mechanical measure is positive and high, even when compared to the return to AFQT In particular, controlling for education, one unit increase in the mechanical test is associated with a 3,58 percent increase in the level of hourly wages. The effect is even bigger than the effect of socio-emotional test scores, although less precise. The effect of the cognitive test on wages is more than twice this value.

So far, the regressions show that mechanical abilities are rewarded by the labor market but imply a different behavior. Those regressions are problematic because 1) schooling choices are endogenous and that must be controlled for if to estimate the returns to unobserved heterogeneity and 2) Test scores are just proxies of abilities and they are influenced by schooling, age and family background variables. The next section presents the model proposed to measure more accurately the effect of mechanical ability.

4.3 Augmented Roy Model with Factor Structure

The model presented in this section is a simplified version of the model presented in chapter 3. We abstract from the selection into occupations to concentrate specifically in the decision to attend four-year college among elegible individuals. Each of the components of the model will be presented in a separate subsection. The model estimated uses one schooling choice (attending a four-year college or not), 3 factors (the three dimensions of ability), 6 cognitive tests, 3 tests on mechanical ability, and 2 tests on socio-emotional abilities.

4.3.1 Model of Schooling Choice

The latent utility of getting education is given by:

$$D = \mathbf{1}[I_i > 0]$$

$$I_i = X_i\beta + \lambda_D^c\theta_{c,i} + \lambda_D^m\theta_{m,i} + \lambda_D^s\theta_{s,i} + e_i \text{ for } i = 1, ...N$$

$$e_i \sim N(0,1)$$

where X_i is a matrix of observed variables that affect schooling, β is the vector of coefficients. $\theta = [\theta_{c,i}, \theta_{m,i}, \theta_{s,i}]$ is the vector of latent abilities where subscript c is used to denote the cognitive ability, subscript m denotes mechanical ability and subscript s denotes socio-emotional ability. $\lambda_D^c, \lambda_D^m, \lambda_D^s$, the vectors of returns to these abilities. These coefficients are referred in the literature as the factor loadings. e_i is the error component that is assumed to be independent of X_D, θ and following a standard normal distribution. Then D denotes a binary variable that takes the value of 1 if the individual chooses to attend a 4-year college and 0 otherwise.[7]

Conditional on X and θ the equations produce a standard discrete choice model with a factor structure. Furthermore, given the set of assumptions exposed, this can be interpreted as the standard probit model.

4.3.2 Model of Hourly Wages

Analogously, the model of earnings can be expressed as a linear function of $X_{w,i}$ and θ in the following way:

$$\ln w_{D,i} = X_{w,i}\beta_{w,D} + \lambda_{w,D}^c\theta_{c,i} + \lambda_{w,D}^m\theta_{m,i} + \lambda_{w,D}^s\theta_{s,i} + e_{w,D,i}$$

$$e_{w,D,i} \sim N(0,1)$$

[7]Through all the exposition the indicator function will be used, $\mathbf{1}[]$ this function takes a value of one if the condition inside the parentheses is satisfied.

for $D = \{0, 1\}$.

4.3.3 Model of Test Scores: Measurement System

Motivated for the findings of the Exploratory Factor Analysis performed in Section 3 the model of test scores allow each measurement to be a function of the corresponding latent ability. For the mechanical tests we allow them to be a function of both cognitive and mechanical latent factors.

In this context, the model for the cognitive measure C_j is:

$$C_{j,i} = X_{C_j,i}\beta_{C_j} + \lambda^c_{C_j}\theta_{c,i} + e_{C_j,i}$$

for $j = \{1, ..., 6\}$.

The model for the mechanical measure M_l is:

$$M_{k,i} = X_{M_k,i}\beta_{M_k} + \lambda^c_{M_k}\theta_{c,i} + \lambda^m_{M_k}\theta_{m,i} + e_{M_k,i}$$

for $k = \{1, ..., 3\}$.

And the model for the socio-emotional measure S_l is:

$$S_{l,i} = X_{S_l,i}\beta_{S_l} + \lambda^s_{S_l}\theta_{s,i} + e_{S_l,i}$$

for $l = \{1, 2\}$.

Finally, all error terms $\{e_i, e_{w,D,i}, e_{C_1,i}, ..., e_{C_6,i}, e_{M_1,i}, ..., e_{M_3,i}, e_{S_1,i}, e_{S_2,i}\}$ for $D =$

$\{0, 1\}$, $j = \{1, ..., 6\}$, $k = \{1, ..., 3\}$ are mutually independent, independent of the factors and independent of all observable characteristics. This independence is essential to the model since it implies that all the correlation in observed choices and measurements is captured by latent unobserved factors.

4.3.4 Latent Factors

The observed level of these latent factors may be the result of some combination of inherited ability, the quality of the family environment in which individuals were raised, cultural differences, etc. These factors are assumed to be fixed by the time the individual is choosing the level of education and thus, by the time the labor and behavioral outcomes considered in this document are determined. In addition, the factors are assumed to be known by the individual but unknown to the researcher. Following standard conventions it is assumed that cognitive and mechanical factors are independent to the Socio-emotional factor while cognitive and mechanical can be correlated.

A mixture of normals is used to model the distribution of the latent abilities. This distribution is selected because as Ferguson (1983) proved, a mixture of normals can approximate any distribution and we want to impose the minimum number of restrictions on the distribution of these unobserved components.

In this case, we use mixtures of two normal distributions (i.e., $K = J = L = 2$) and assume $E[\theta_c] = E[\theta_m] = E[\theta_s] = 0$. Finally, we impose $(\theta_c, \theta_m) \perp \theta_s$. For more details on this and the identification strategy refer to Appendix 2.

4.3.5 Estimation Strategy

Let $T_i = \{C_{1i}, ..., C_{6,i}, M_{1i}, ..., M_{3,i}, S_{1i}, S_{2,i}\}$ be the vector of test scores for individual i, $X_{T,i} = \{X_{C,i}, X_{M,i}, X_{S,i}\}$ and $\theta = [\theta_c, \theta_m, \theta_s]$ the vector of the latent factors and δ the vector of all the parameters of the model. Thus, our likelihood function is:

$$L(\delta|X) = \prod_{i=1}^{N} f(D_i, \ln w_{D,i}, T_i | X_i, X_{w,i}, X_{T,i})$$

Given that conditional on unobserved endowments all the errors are mutually independent, our likelihood can also be expressed as:

$$L(\delta|X) = \prod_{i=1}^{N} \int_{\Theta} f(D_i, \ln w_{D,i}, T_i | X_i, X_{w,i}, X_{T,i}, \theta) dF(\theta)$$

The model is estimated using MCMC techniques. The use of Bayesian methods in this paper is merely computational to avoid the computation of a high order integral. In consequence, the interest is primarily on the mean of the posterior distribution. Thus, it is viewed from a classical perspective and interpreted as an estimator that has the same asymptotic sampling distribution as the maximum likelihood estimator. See Hansen et al. (2004) and Heckman et al. (2006) for more details.

4.4 Results

We first compare the distribution of the estimated factors with the observed distribution of the measurements. Then we summarize the main results of the model. Anticipating our main findings, we confirm the results obtained from the reduced-form estimates: Mechanical ability reduces the probability of seeking a professional degree and at the same time, it is positively rewarded in the labor market. We use simulations from our model to explore the implications of being low in the standard types of ability but having high levels of mechanical ability in terms of schooling choices and earnings.The model fits the data on wages and college attendance. Goodness of fit test are passed and the three factors are needed in order to fit the data on wages.[8]

4.4.1 Observed Test Scores and Estimated Abilities

This paper treats observed cognitive, socio-emotional, and mechanical test scores as the outcomes of a process that has as inputs family background, schooling at the time of the test and unobserved abilities. Here we present the estimated parameters of the distribution of unobserved abilities as well as the fraction of the variance of observed test scores that can be explained with and without the inclusion of unobserved abilities.

Table 4.5 presents the coefficients on unobserved abilities for each of the tests used. For identification purposes, one loading for each unobserved ability is set to

[8] See Tables 4.15 and 4.16 in Appendix 2.

one. The remaining loadings are interpreted in relation to the loading set as the numeraire (for details see Carneiro et al., 2003, and Appendix 2). The selected numeraires are arithmetic reasoning, mechanical comprehension and the Rosenberg self-esteem scale for cognitive, mechanical and socio-emotional abilities, respectively.

To analyze the relative importance of each dimension of ability in explaining test scores, Figure 4.5 presents the variance decomposition of the measurement system. The results show the contribution of observed variables, latent abilities and error terms as determinants of the variance of each test score.

The variance decomposition illustrates the large size of the unexplained component and highlights the consequences of using observed test scores as proxies for unobserved abilities. The contribution of observed variables to the variance of the test scores is never more than 20 percent. After controlling for the latent variables, the error term is still large but we are able to explain a much higher percentage of the total variance, between 34 and 65 percent. The one exception is the Rotter Scale, where we are only able to explain 11 percent of the variance.

We allow both cognitive and mechanical abilities to influence mechanical test scores. While cognitive ability has lower loadings compared to mechanical ability (see Table 4.5), both abilities are important determinants of the variance in the observed scores.[9]

[9] In a model where mechanical test scores are explained by observed variables and only the cognitive factor, the fraction of the variance explained reduces to a third or two thirds of the fraction that is explained jointly by the two factors.

Distribution of Abilities

Observed test scores and unobserved abilities are different. In this section we use the estimated parameters for the distribution of each ability to estimate the distribution of cognitive, socio-emotional, and mechanical abilities. We show that the distribution of abilities is very different to the distribution of test scores. For mechanical ability, accounting for this difference is especially important as the implied sorting into schooling is completely different when using observed test scores. The mean and standard deviation of the simulated distribution for each ability are displayed in Table 4.6.

Figures 4.6 and 4.7 present the marginal cumulative distribution function of the estimated factor by schooling for the cognitive and socio-emotional, and mechanical abilities respectively. For cognitive and socio-emotional ability (figure 4.6) the cummulative distribution function (cdf) of the ability for people that attended college stochastically dominates the cdf curve for those who did not. Although the distributions are different, the sorting into schooling is similar. In particular, for both observed test scores and unobserved abilities, the cdf for people with high education stochastically dominates the cdf curve for people with low schooling (see figure 4.3).

However, for mechanical ability the relationship is reversed. The distribution of the estimated factor implies that people with high levels of mechanical ability choose low education. The cummulative distribution function (cdf) of the estimated ability for people that chose to attend four-year college is stochastically dominated by the cdf curve for those that did not attend college (see figure 4.7). As a consequence, for mechanical ability, the sorting implied by the estimated factor and the observed test scores is completely different in terms of schooling.

The sorting implied by the estimated factor explains why after controlling for the three scores in the reduced-form estimations, the coefficient of the composite mechanical test in the probit of college attendance changed its sign (see section 4.2).

4.4.2 Effect of Abilities on Schooling Choice and Hourly Wages

Figures 4.8 to 4.13 present the main results of the model in terms of the outcomes of interest: a) the choice of attending a 4-year college and b) log hourly wages. We present two types of figures: joint distributions of the outcome variables by deciles of the factors and marginal effects of each factor on the outcomes of interest integrating out the effect of the other factors.

Figures 4.8 and 4.9 present the joint distribution of the probability of attending a 4-year college reported by deciles of cognitive and mechanical and by the deciles of socio-emotional and mechanical, respectively.

In the first case, the opposite effects of the abilities are evident but the positive effect of cognitive is always stronger. As en exercise we can move along the distributions and compare the effect of increasing one decile on both cognitive and mechanical on the probability of going to college. Given that cognitive has a positive effect and mechanical a negative effect this exercise will show which effect prevails.

Starting at the lowest extreme of both distributions (first decile of both cognitive and mechanical) and moving to the next decile of the distributions of both cognitive and mechanical abilities the estimated probability of going to college always increases.

A similar exercise on the distributions of socio-emotional and mechanical shows a very flat slope. This is a consequence of the correlation of mechanical and cognitive ability and the opposite effects of mechanical and socio-emotional ability (see Figure 4.9).

The marginal effect of cognitive ability integrating out the effect of mechanical is presented in panel a of Figure 4.10 while panel b and c present the analogous for socio-emotional and mechanical ability, respectively.

Table 4.7 presents the effect on college attendance associated with a one standard deviation increase in each of the factors. According to the estimates, one standard deviation increase in cognitive ability is associated with an increase of 19.3 percentage points in the probability of attending 4-year college, the same increase in socio-emotional ability is associated with a 2.7 increase in the probability while one standard deviation increase in mechanical ability decreases the probability in 7.5 percentage points.

Figures 4.11 and 4.12 present the total effect of ability on log wages, including the direct effect of ability on log wages holding schooling constant, the effect of ability on the decision to attend college and the implied effect of attending or not college on log wages. The effect is positive for all three dimensions of ability.

The marginal effect of mechanical ability is considerable small compared with the effect of cognitive and also with the effect of socio-emotional ability (Figure 4.13). In fact, a one standard deviation increase in cognitive ability is associated with 9.8 percent increase in log hourly wages and 3.9 for socio-emotional ability while the average estimated effect of mechanical is 1.4 percent (see the last row of table 4.8).

The story changes when analyzing the returns to ability by college attendance. In the case of not attending a four-year college the returns to cognitive and mechanical ability are very close, 4.7 and 4.4 precent, respectively. While in the case of attending college the returns to cognitive ability are 10.8 percent compared to the -3.1 percent in the case of mechanical ability. For socioemotional ability the difference in the returns is smaller although the returns are higher in the scenario of college attendance.

4.5 Discussion

In this section we analyze the implications of our results in terms of the wage gains asociated with college attendance for individuals with diferent ability profiles. In particular, we are interested in understanding the implications of having low levels of cognitive and socio-emotional ability but high levels of mechanical ability.

Using the estimates from the model we compute the difference between the mean of hourly wages conditional on the schooling choice and the respective counterfactual wage.

$$E[Y_0|D=0] - E[Y_1|D=0] = E[Y_0 - Y_1|D=0]$$

$$E[Y_1|D=1] - E[Y_0|D=1] = E[Y_1 - Y_0|D=1]$$

On average the mean of hourly wages conditional on college attendance is 10 percent higher than the respective counterfactual (i.e., the wage that would have been received if the individual had decided not attending to college). In contrast, conditioning on not attending college the mean of hourly wages is 3.8 percent lower than the mean of the counterfactual. These results would suggest that college is associated with higher wages even for individuals that, given their observable characteristics and latent abilities, decided not attending college.

But this average result does not hold for all individuals, particularly given the special behavior implied by mechanical ability. With this in mind, we investigate the gains of not attending college conditional on the decision of not attending, $E[Y_0 - Y_1|D=0]$, for different ability profiles.

Table 4.10 presents the results using the quintiles of the distribution of ability to define specific profiles. The columns correspond to the bottom, middle and top quintiles of mechanical ability and the rows present four extreme ability profiles defined as a combination of different levels of cognitive and socio-emotional ability. The first row corresponds to the low ability profile, which means an individual in the lowest quintile of both cognitive and socio-emotional; the second row displays the low cognitive high socio-emotional profile (in the first quintile of the distribution of cognitive ability and 5th quintile of the distribution of socio-emotional ability; row three presents the opposite case, high cognitive and low socio-emotional; row four presents the high ability type (highest quintile of the distribution of both cognitive and socio-emotional ability).

Given the high return to college education most of the cells in the table are positive. But for individuals in the highest quintile of mechanical ability, the conditional mean of hourly wages is higher than the alternative when the other two abilities are in the bottom of the distribution and also when cognitive is low and socio-emotional is high. This suggests that individuals with very high levels of mechanical ability but low levels of cognitive ability not going to college is associated with the highest expected hourly wage.[10]

Finally, we analyze the composition of the population that benefits from not going to college (22 percent of the population). Nearly 65 percent of those who benefit are individuals above the median of the distribution of mechanical ability summing up to 14 percent of the total population (See Figure 4.14).

Although the absolute percentages are useful, it is important to take into account that the amount of population in each specific profile varies. More specifically, the positive correlation between mechanical and cognitive ability would necessarily imply that the amount of population with high levels of both abilities is always higher that the amount of population with low levels of one and high levels of the

[10]According to the estimated distributions of abilities close to 3.5 percent of the population are low cognitive, low socio-emotional and high mechanical ability.

other. Figure 4.15 shows that almost 40 percent of the individuals with low cognitive, low socio-emotional and high mechanical ability benefits from not going to college. That percentage decreases pregressively for the low cognitive-high socio-emotional, the high cognitive low socioemotional and the high cognitive and high socioemotional combinations. In consequence, nearly 28 percent of the individuals with high mechanical ability and 15 percent of the individuals with low mechanical ability would obtain a positive difference between the observed hourly wage and the counterfactual wage conditional on the decision of not attending college.

4.6 Conclusions

This paper investigates the role of mechanical ability in explaining schooling decisions and labor market outcomes. We show that this dimension of ability is positively rewarded by the labor market, but in contrast to the conventional facets of ability, it predicts the choice of lower levels of education. In particular, controlling for cognitive and socio-emotional aspects, mechanical ability reduces the likelihood of attending a four-year college. As a consequence, mechanical ability comes to enrich the set of factors explaining the observed disparities in schooling decisions and labor market outcomes.

But we do more than simply expand the range of empirically relevant dimensions of abilities. In fact, by including mechanical ability in the analysis we alter the dichotomous paradigm of low and high ability individuals in the context of the previously accepted symmetry of the impact of abilities on schooling decisions and labor market productivity.

Our results suggest a new framework where individuals with low levels of cognitive and socio-emotional ability, may have high mechanical ability and greatly benefit from it. More precisely, we find that despite the high return associated with college attendance, these individuals could expect higher wages by choosing not to attend a four-year college. This conclusion is a direct result of the high returns to mechanical ability in jobs not requiring a four-year college degree which contrast with the negative returns to mechanical ability in jobs requiring it.

The results from our empirical model highlight the importance of moving beyond the "one-size-fits-all" college discourse and explore alternative pathways to successful careers for individuals with a different profile of skills. This message is particularly relevant in a nation where less than half of the students attempting to get a bachelor's degree actually get one and where completion rates are below 20 percent for students who score low in standardized achievement tests during high school.[11] Accepting the multidimensional nature of ability must be accompanied by the implementation of inclusive human capital development strategies with more than one pathway to success.

As a final note, this article leaves some important areas for extensions and future research. First, the analysis of wage growth and the comparison between initial versus late returns to skill. There are many reasons to expect a lower wage gradient for skills in early career spans and the current model does not account for that. Second, it would be interesting to extend the model to analyze gender and

[11] NCES (2013) and Rosenbaum et al., 2010.

race disparities.

4.7 Tables and Figures

Table 4.1: Correlation of the Technical Composites of the ASVAB with Tests Used to Create AFQT and a Composite Measure of Socio-emotional

	Auto	Mech	Elect	AFQT	Arith	Coding	Math	Word	Parag	Num	SocioE
Auto	1.00										
Mechanical. C	0.68	1.00									
Electronics	0.69	0.70	1.00								
AFQT	0.49	0.64	0.66	1.00							
Arithmetic K.	0.45	0.62	0.59	0.87	1.00						
Coding S.	0.32	0.42	0.40	0.76	0.54	1.00					
Math	0.31	0.53	0.51	0.85	0.78	0.54	1.00				
Word K.	0.56	0.61	0.71	0.83	0.66	0.50	0.62	1.00			
Paragraph C.	0.48	0.58	0.62	0.84	0.67	0.53	0.63	0.77	1.00		
Numerical S.	0.31	0.41	0.42	0.81	0.62	0.67	0.61	0.55	0.57	1.00	
SocioEmot.	0.23	0.25	0.26	0.31	0.26	0.21	0.23	0.33	0.28	0.25	1.00

Note: AFQT is the cognitive measure, it represents the standardized average over the ASVAB score in six of the ten components: math knowledge, arithmetic reasoning, word knowledge, paragraph comprehension, numerical speed and coding speed. Socio-emotional is the standardized average of the scores for the Rotter and Rosenberg tests.

Table 4.2: Summary statistics

Variable	Mean	(Std. Dev.)	Min.	Max.	N
LogHourly wage 25-30	2.812	(0.41)	0.628	4.053	1385
Attended 4yrcollege by age 25	0.321	(0.467)	0	1	1466
Urban residence at age 25	0.704	(0.457)	0	1	1355
Northeast residence at age 25	0.175	(0.38)	0	1	1466
Northcentral residence at age 25	0.33	(0.47)	0	1	1466
South residence at age 25	0.255	(0.436)	0	1	1466
West residence at age 25	0.158	(0.365)	0	1	1466
Cohort1 (Born 57-58)	0.126	(0.332)	0	1	1466
Cohort2 (Born 59-60)	0.19	(0.392)	0	1	1466
Cohort3 (Born 61-62)	0.334	(0.472)	0	1	1466
Cohort4 (Born 63-64)	0.351	(0.477)	0	1	1466
Family Income in 1979 (thousands)	21.878	(11.849)	0	75.001	1466
Broken home at age 14	0.193	(0.395)	0	1	1463
Number of siblings 1979	2.934	(1.887)	0	13	1466
Mother's highest grade completed	11.442	(3.196)	0	20	1466
Father's highest grade completed	11.535	(3.985)	0	20	1466
Living in urban area at age 14	0.726	(0.446)	0	1	1466
Living in the south at age 14	0.248	(0.432)	0	1	1466
Education at the time of the test	11.22	(1.011)	6	12	1466
AFQT	0	(1)	-3.328	2.007	1466
Mechanical	0	(1)	-3.348	1.985	1466
SocioEmotional	0	(1)	-2.718	2.452	1466

Notes: AFQT is an average of standarized scores for arithmetic reasoning, word knowledge, paragraph comprehension, mathematics knowledge, numerical operations and coding speed sections of the ASVAB. Socio-emotional is an average of the scores in two tests: Rotter Locus of Control Scale and Rosenberg Self-Esteem Scale. Mechanical is an average of standarized scores for auto and shop information, mechanical comprehension and electronics information sections of the ASVAB.

Table 4.3: Schooling Choice: Probit of College Attendance

	(1)	(2)	(3)
AFQT	0.175***		0.206***
	(0.0154)		(0.0177)
Socio-emotional	0.0161	0.0411***	0.0188
	(0.0133)	(0.0133)	(0.0134)
Mechanical		0.0351**	-0.0623***
		(0.0139)	(0.0163)
Observations	1466	1466	1466
Pseudo R^2	0.261	0.176	0.271

Marginal effects; Standard errors in parentheses

(d) for discrete change of dummy variable from 0 to 1

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Sample: males between 25-30 years old,not attending school and up to high school complete by the time of the test. * Marginal effects at the mean. All regressions include family background controls, cohort dummies and geographical controls for region and urban residence at the age of 14

Table 4.4: Log Hourly Wages: OLS

	(1)	(2)	(3)
College	0.142***	0.214***	0.151***
	(0.0378)	(0.0353)	(0.0380)
AFQT	0.106***		0.0857***
	(0.0167)		(0.0200)
Socio-emotional	0.0359**	0.0433***	0.0338**
	(0.0158)	(0.0158)	(0.0158)
Mechanical		0.0811***	0.0358*
		(0.0161)	(0.0192)
Observations	1355	1355	1355
R^2	0.115	0.104	0.117

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Sample: males between 25-30 years old,not attending school and up to high school complete by the time of the test. College is dummy variable for college degree or more. All regressions include cohort dummies as well as geographical controls for region and urban residence at age 25.

Table 4.5: Loadings on Test Scores

	Cognitive		Mechanical		Socio-emotional
Auto	0.55	***	1.32	***	
Electronics	0.43	***	0.88	***	
Mech. C	0.38	***	1.00		
Arithmetic K.	1.06	***			
Math	1.00				
Word K.	0.96	***			
Paragraph C.	0.97	***			
Numerical S.	0.79	***			
Coding S.	0.73	***			
Rotter					0.26***
Rosenberg					1.00

All regressions include family background controls (mother's and father's education, number of siblings, a dummy for broken family at age 14, family income in 1979), cohort dummies and geographical controls for region and urban residence at the age of 14.

Table 4.6: Simulated Parameters of the Distribution of Ability

	Mean	SD	Covar(θ^c,θ^i)	Correlation(θ^c, θ^i)
θ^c	-0.001	0.73	0.53	1
θ^m	0.000	0.58	0.21	0.52
θ^s	-0.001	0.89	0	0

Note: Results simulated from the estimates of the model and our NLSY79 sample

Table 4.7: Estimated Marginal Effects: College Attendance

	Cognitive	**Mechanical**	**Socio-emotional**
College Decision	0.229	-0.095	0.024
	(0.002)***	(0.001) ***	(0.0000) ***

Note: Standard errors in parenthesis. College Decision equation includes family . background controls, cohort dummies and geographical controls for region and urban residence at the age of 14.

Table 4.8: Estimated Marginal Effects: Log of Hourly Wages

	Cognitive	**Mechanical**	**Socio-emotional**
College=0 (w0)	0.047	0.044	0.033
	(0.002)***	(0.001)***	(0.000)***
College=1 (w1)	0.108	-0.031	0.047
	(0.002)***	(0.001) ***	(0.001) ***
Overall	0.107	0.014	0.041
	(0.000)***	(0.001) ***	(0.001) ***

Note: Standard errors in parenthesis. We control for cohort dummies as well . as geographical controls for region and urban residence at age 25.

Table 4.9: Comparative Advantage

Formula	**Estimate**
$E[Y_1\|D=1] - E[Y_0\|D=1]$	0.102***
$E[Y_0\|D=0] - E[Y_1\|D=0]$	-0.038***

Table 4.10: $E[Y_1 - Y_0|D=0]$ by Quintiles of Mechanical Ability and Different Levels of Cognitive and Socio-emotional Abilities

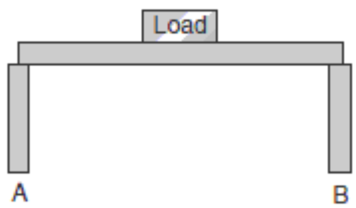
Mechanical	**Quintile 1**		**Quintile 3**		**Quintile 5**	
Low C - Low S	10.4%	***	0.6%		-6.8%	***
Low C - High S	14.5%	***	4.8%	***	-3.9%	**
High C - Low S	24.6%	***	13.1%	***	5.3%	***
High C - High S	25.8%	***	18.0%	***	9.0%	***

Low refers to the first quintile of the distribution of Cognitive (C) or Socio-emotional (S), while High refers to the fifth quintile.

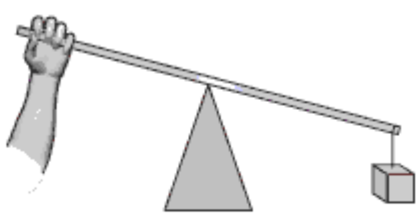
Figure 4.1: Sample question from the mechanical comprehension section

[a]

1. In the diagram, what can you tell about the load on posts A and B?
   (a) Post B carries more weight.
   (b) Post A carries more weight.
   (c) Post A carries no weight.
   (d) The load is equal on posts A and B.



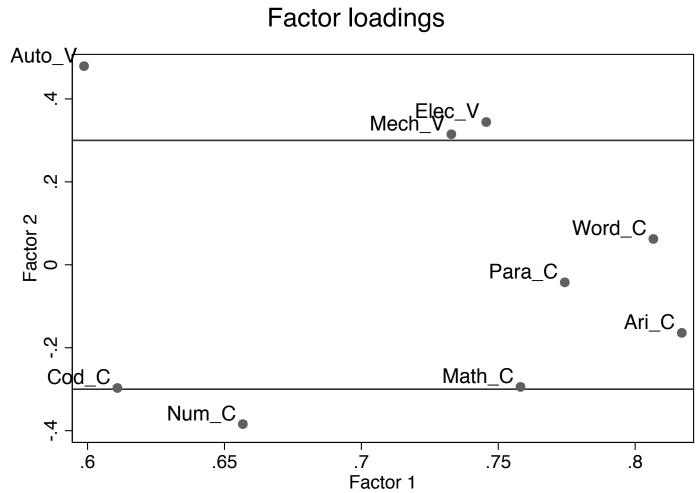
2. The diagram shows a class 1 lever. Which of the following is the same kind of lever?
   (a) A pair of tweezers
   (b) A pair of scissors
   (c) A wheelbarrow
   (d) A pair of tongs



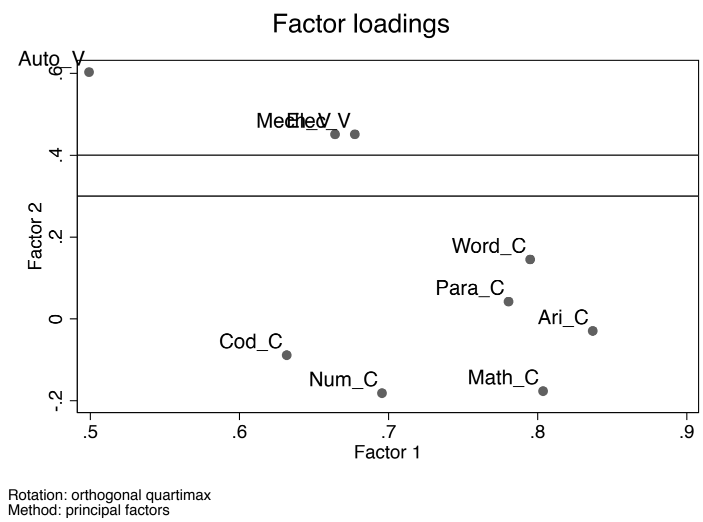
3. Which of the following would feel hottest to the touch if one end were placed in a pot of boiling water?
   (a) A wooden spoon
   (b) A metal fork
   (c) A plastic knife
   (d) A plastic cup

[a] Extracted from http://www.education.com/reference/article/mechanical-comprehension-quiz/

Figure 4.2: Loadings from Factor Analysis-Orthogonal Factors



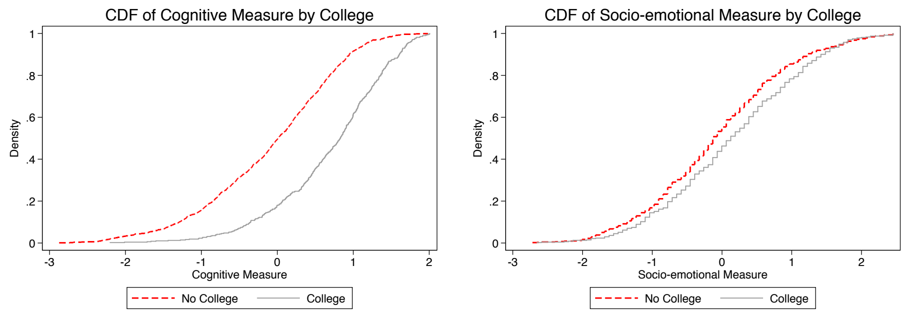
(a) Unrotated



(b) Rotated

Mechanical is computed by using the three first test that appear in the graph: Auto_V (automotive and shop information), Mech_V (mechanical comprehension) and Elec_V (electronics information). The others are used to measure the cognitive component: Ari_C (arithmetic reasoning), Math_C (mathematics knowledge), Word_C (word knowledge) and Para_C (paragraph comprehension) Num_C (numerical operations) and Cod_C (coding speed). All are used to compute AFQT except from Cod_C. In fact, the calculation of AFQT has changed considerably on time. In 1980 it was computed as the raw sum of arithmetic reasoning, word knowledge, paragraph comprehension and 1/2 numerical operations. After 1989 numerical operations was removed and mathematics knowledge was included.

Figure 4.3: Measurement of Cognitive and Socio-emotional Ability



(a) Cognitive (b) Socio-emotional

Figure 4.4: Measurement of Mechanical Ability

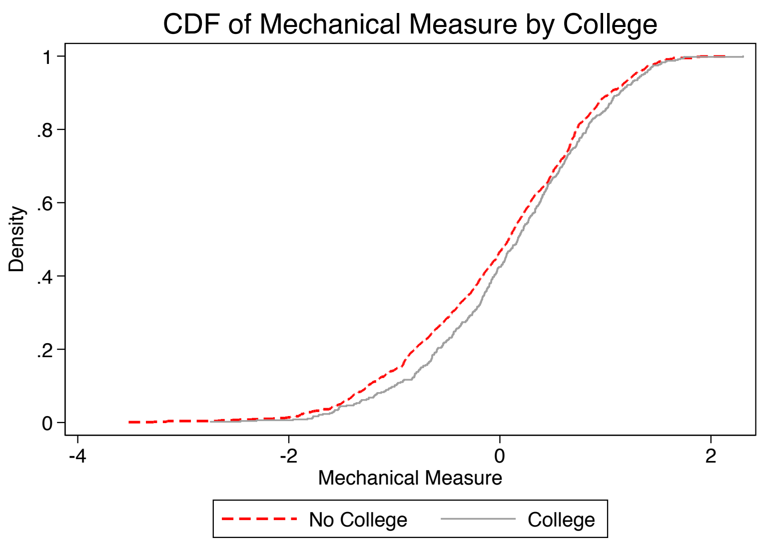


Figure 4.5: Variance Decomposition

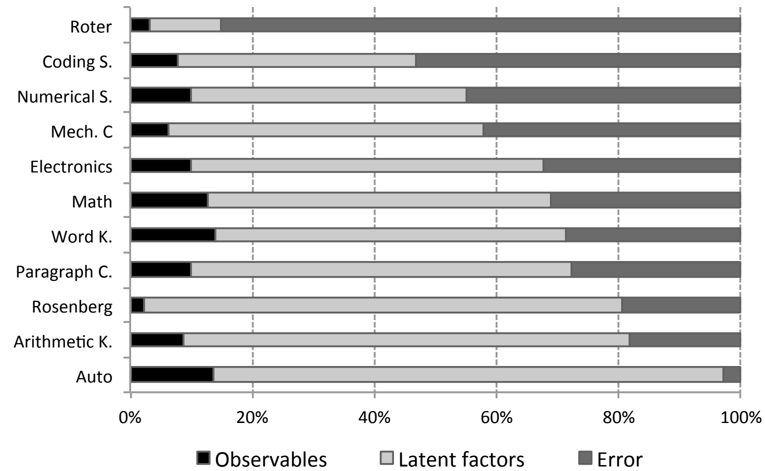
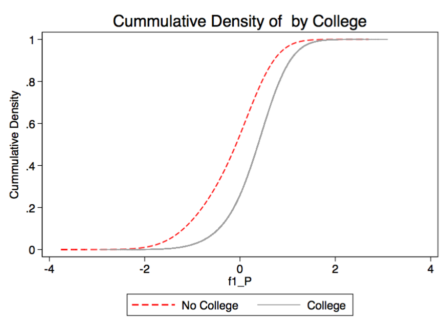
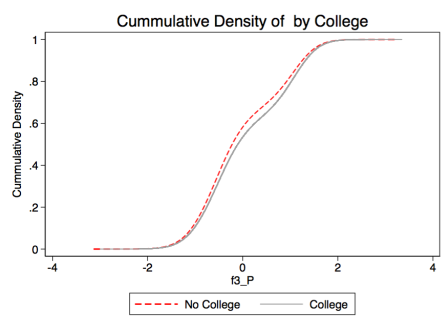


Figure 4.6: Marginal CDF: Cognitive and Socio-emotional Ability



(a) Cognitive



(b) Socio-emotional

Figure 4.7: Marginal CDF: Mechanical Ability

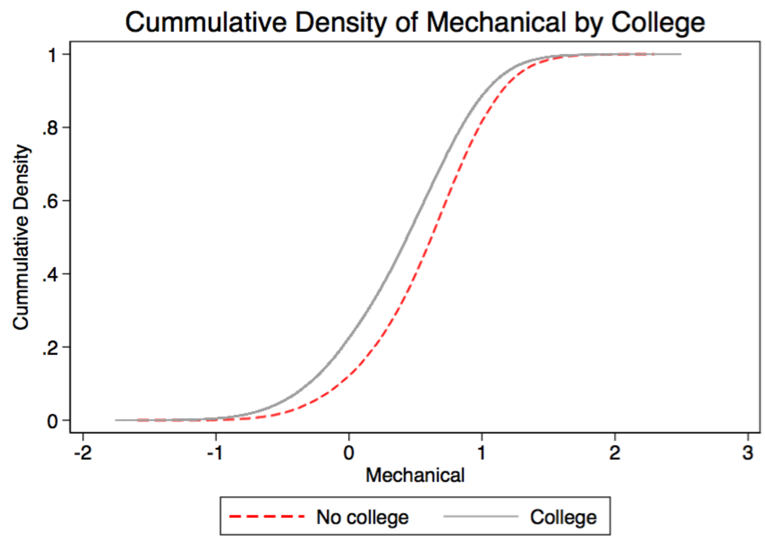
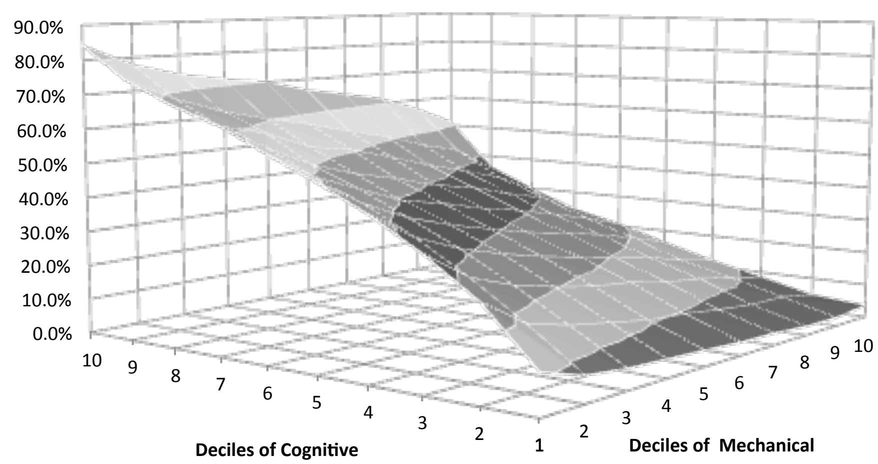
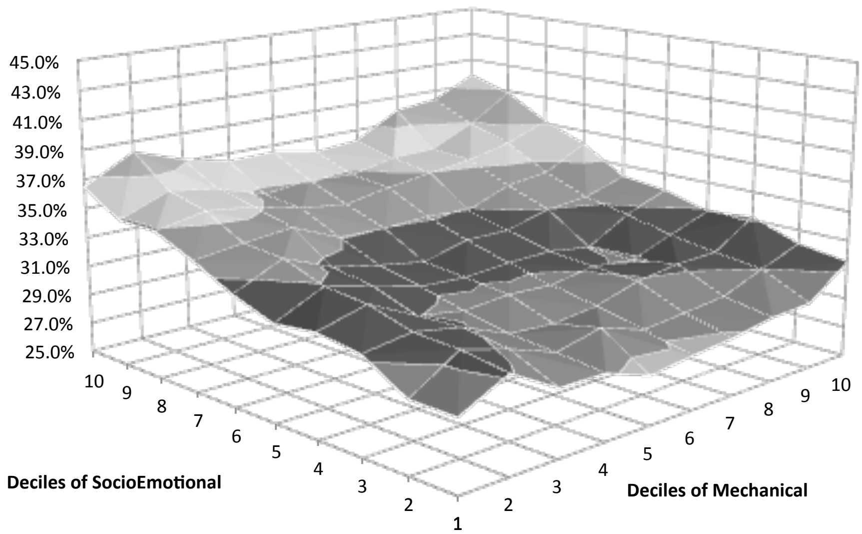


Figure 4.8: Joint Distribution of College Attendance Decision by Deciles of Cognitive and Mechanical Factors



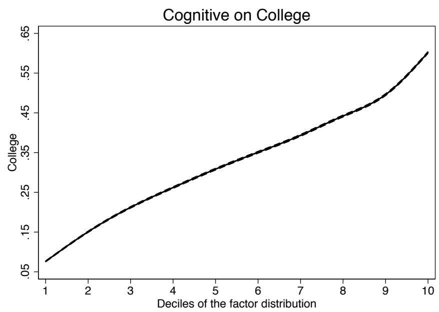
Note: The data are simulated from the estimates of the model and our NLSY79 sample. In the figure we plot $P_{i,j} = \int (Pr(D = 1|\theta_c = d_i, \theta_m = d_j))\, dF\theta_s$ for $d_i = 1, ..10$ and $d_j = 1, ..10$

Figure 4.9: Joint Distribution of College Attendance Decision by Deciles of Socio-emotional and Mechanical Factors

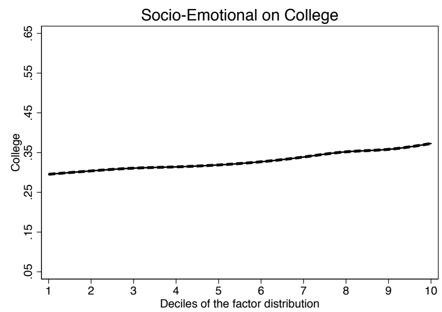


Note: The data are simulated from the estimates of the model and our NLSY79 sample. In the figure we plot $P_{i,j} = \int (Pr(D = 1|\theta_m = d_i, \theta_s = d_j))\, dF\theta_c$ for $d_i = 1, ..10$ and $d_j = 1, ..10$

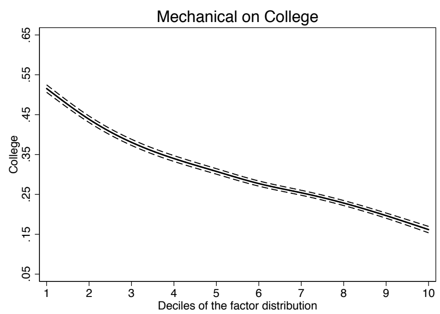
Figure 4.10: Marginal Effect of Ability on College Attendance



(a) Cognitive



(b) Socio-emotional



(c) Mechanical

Note: The data are simulated from the estimates of the model and our NLSY79 sample.

Figure 4.11: Average of Log Wage by Deciles of Cognitive and Mechanical Factors

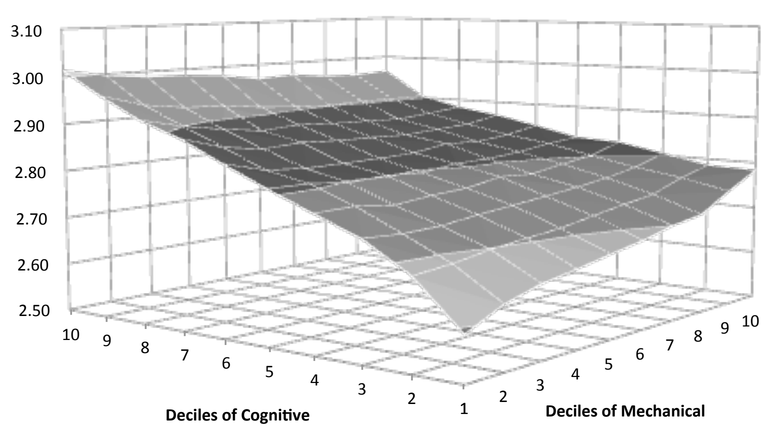


Figure 4.12: Average of Log Wage by Deciles of Socio-emotional and Mechanical Factors

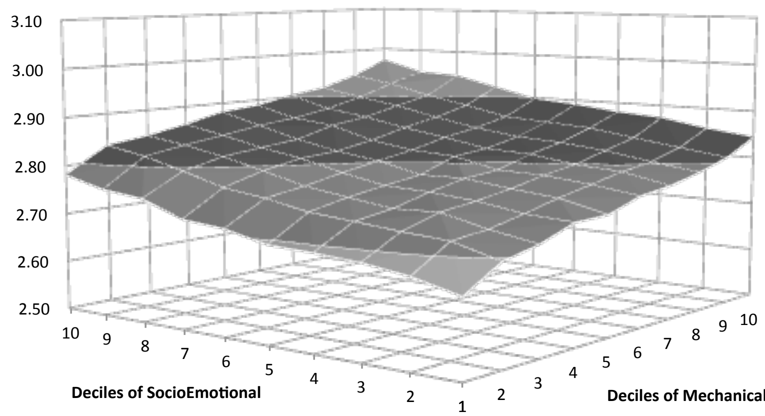
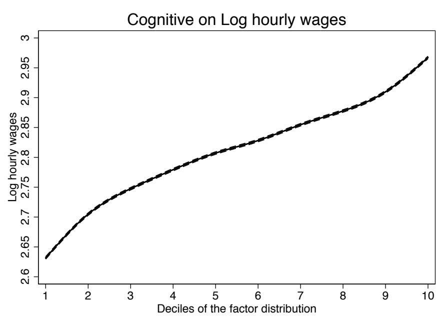
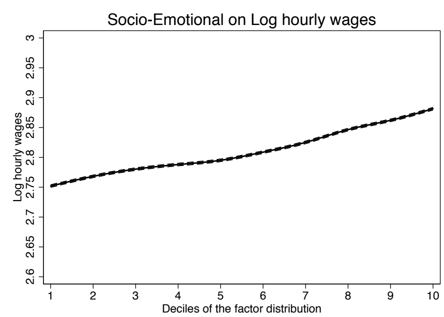


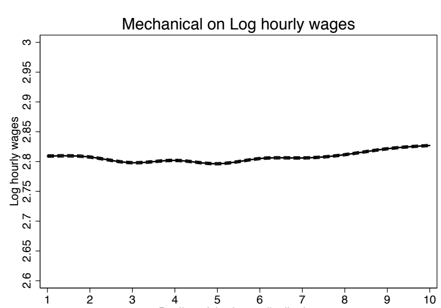
Figure 4.13: Marginal Effect of Ability on Log Hourly Wages



(a) Cognitive

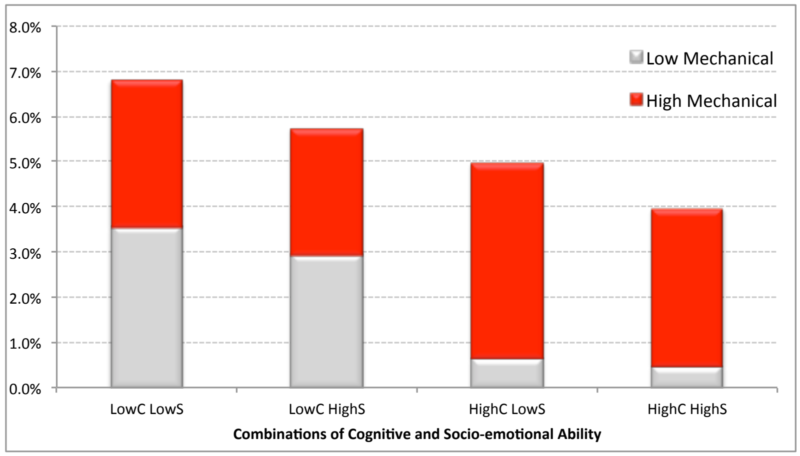


(b) Socio-emotional



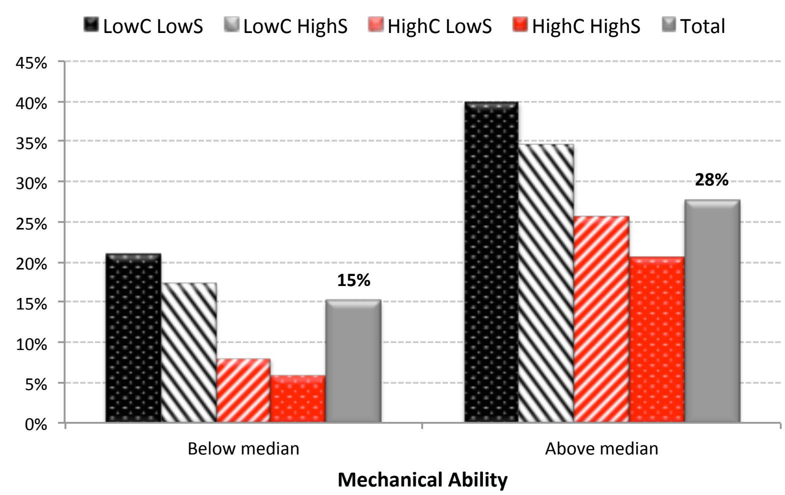
(c) Mechanical

Figure 4.14: Profile Composition of the Individuals that Benefit from not Attending College



Note: The data are simulated from the estimates of the model and our NLSY79 sample.

Figure 4.15: Who Benefits from not Attending College?



Note: The data are simulated from the estimates of the model and our NLSY79 sample. Figure presents the percentage of people that benefits from not attending college in each category.

Table 4.11: Estimates of the Model: Measurement Equations

	cons	Sibl	Med	Fed	FamY	urban	south	coh1	coh2	coh3	hgtest	c	m	s
Auto	-2.64	-0.02	0.01	0.01	0.00	-0.16	-0.19	0.53	0.34	0.07	0.23	0.55	1.32	
SE	0.39	0.01	0.01	0.01	0.00	0.05	0.06	0.10	0.09	0.08	0.04	0.04	0.08	
Elec	-2.93	-0.05	0.01	0.02	0.00	-0.07	-0.17	0.20	0.04	-0.09	0.25	0.43	0.88	
SE	0.39	0.01	0.01	0.01	0.00	0.05	0.06	0.10	0.09	0.08	0.04	0.04	0.05	
Mech	-2.94	-0.01	0.02	0.01	0.00	-0.15	-0.15	-0.06	-0.17	-0.18	0.25	0.38	1.00	
SE	0.40	0.01	0.01	0.01	0.00	0.05	0.06	0.10	0.09	0.08	0.04	0.04	0.00	
Arith	-3.40	0.00	0.03	0.02	0.00	-0.02	-0.19	-0.30	-0.44	-0.34	0.27	1.06		
SE	0.39	0.01	0.01	0.01	0.00	0.05	0.05	0.09	0.09	0.08	0.04	0.03		
Math	-2.83	-0.02	0.02	0.04	0.01	-0.01	-0.19	-0.60	-0.62	-0.25	0.21	1.00		
SE	0.37	0.01	0.01	0.01	0.00	0.05	0.06	0.09	0.09	0.08	0.04	0.00		
Word	-3.80	-0.05	0.03	0.03	0.00	-0.04	-0.13	-0.10	-0.30	-0.34	0.30	0.96		
SE	0.38	0.01	0.01	0.01	0.00	0.05	0.06	0.09	0.09	0.08	0.04	0.03		
Para	-3.51	-0.02	0.02	0.04	0.00	-0.05	-0.06	-0.31	-0.39	-0.29	0.28	0.97		
SE	0.38	0.01	0.01	0.01	0.00	0.05	0.06	0.10	0.09	0.08	0.04	0.04		
Num	-3.49	-0.02	0.02	0.02	0.01	-0.01	-0.14	-0.24	-0.41	-0.24	0.27	0.79		
SE	0.37	0.01	0.01	0.01	0.00	0.06	0.06	0.10	0.09	0.08	0.04	0.03		
Cod	-2.98	-0.01	0.01	0.02	0.01	0.01	-0.18	-0.14	-0.13	-0.19	0.23	0.73		
SE	0.38	0.01	0.01	0.01	0.00	0.05	0.06	0.10	0.10	0.08	0.04	0.04		
Rotter	-1.93	0.00	0.00	0.01	0.00	0.00	-0.02	0.08	-0.04	-0.08	0.15			0.26
SE	0.40	0.01	0.01	0.01	0.00	0.06	0.06	0.11	0.10	0.08	0.04			0.03
Rosen	-0.82	-0.02	0.01	0.01	0.00	0.00	0.00	0.18	0.18	0.16	0.05			1.00
SE	0.38	0.01	0.01	0.01	0.00	0.05	0.05	0.10	0.09	0.08	0.04			0.00

Note: This table presents estimates of the model. Using data from the NLSY79, white males between 25-30 years old. Since the model is estimated using Bayesian methods, they represent the mean estimates over 1,000 iterations after discarting the first 30,000. The computation of standard errors is explained in appendix B. cons is the constant, Sib is the number of siblings in 1979, Med is the mother's highest grade completed at age 17, Fed is the father's highest grade completed at age 17, FamY is the family income in 1979 in thousands, urban is a dummy variable for living in an urban area at age 14, south is a dummy variable for living in the south at age 14, Coh1 refers to the first cohort (born 57-58), Coh2 refers to the second (born 59-60), Coh3 refers to the last cohort of individuals, those that were born between 61-62, hgtest is the highest grade attended by the time the test was presented and c, m, s refers to the cognitive, mechanical and socio-emotional factors respectively. The first three rows refer to the scores in the technical composites of the ASVAB, the next six scores are the tests used to capture cognitive ability and the last two rows are the socio-emotional test scores.

4.8 Appendix

4.8.1 Appendix 1: Additional Tables and Figures

4.8.2 Appendix 2: Goodness of Fit and Comparison with a Two-Factor Model

In this appendix we present evidence on the goodness of fit for hourly wages and college attendance. Also, we demonstrate that our proposed three-factor model does

Table 4.12: Estimates of the Model: College Decision Model

Pr(Attending college)	Coefficient	SE
Constant	-2.02	0.25
Number of siblings	-0.06	0.03
Mother's highest grade completed	0.05	0.02
Father's highest grade completed	0.09	0.01
Family Income 1979 (thousands)	0.01	0.00
Living in urban area at age 14	0.12	0.11
Living in the south at age 14	0.05	0.11
Cohort1 (Born 57-58)	-1.42	0.19
Cohort2 (Born 59-60)	-1.11	0.14
Cohort3 (Born 61-62)	-0.36	0.11
Cognitive	1.22	0.09
Mechanical	-0.74	0.12
Socio-emotional	0.11	0.05

Note: This table presents estimates of the model. Using data from the NLSY79, white males between 25-30 years old. Since the model is estimated using Bayesian methods, they represent the mean estimates over 1,000 iterations after discarding the first 30,000. The computation of standard errors is explained in appendix 3.

Table 4.13: Estimates of the Model: Log of Hourly Wage

	No college	SE	College	SE
Constant	2.83	0.05	2.91	0.06
Northeast residence	0.02	0.04	0.22	0.06
Northcentral residence	-0.11	0.04	0.01	0.06
South residence	-0.13	0.04	0.03	0.06
Cohort2 (Born 59-60)	0.01	0.03	-0.02	0.07
Cohort3 (Born 61-62)	-0.03	0.03	-0.02	0.04
Local Unemployment rate	0.08	0.46	-1.50	0.65
Cognitive	0.06	0.02	0.15	0.04
Mechanical	0.08	0.03	-0.05	0.05
Socio-emotional	0.04	0.02	0.05	0.02

Note: This table presents estimates of the model. Using data from the NLSY79, white males between 25-30 years old. Since the model is estimated using Bayesian methods, they represent the mean estimates over 1,000 iterations after discarding the first 30,000. The computation of standard errors is explained in appendix 3.

Table 4.14: Parameters of the Distribution of Unobserved Abilities

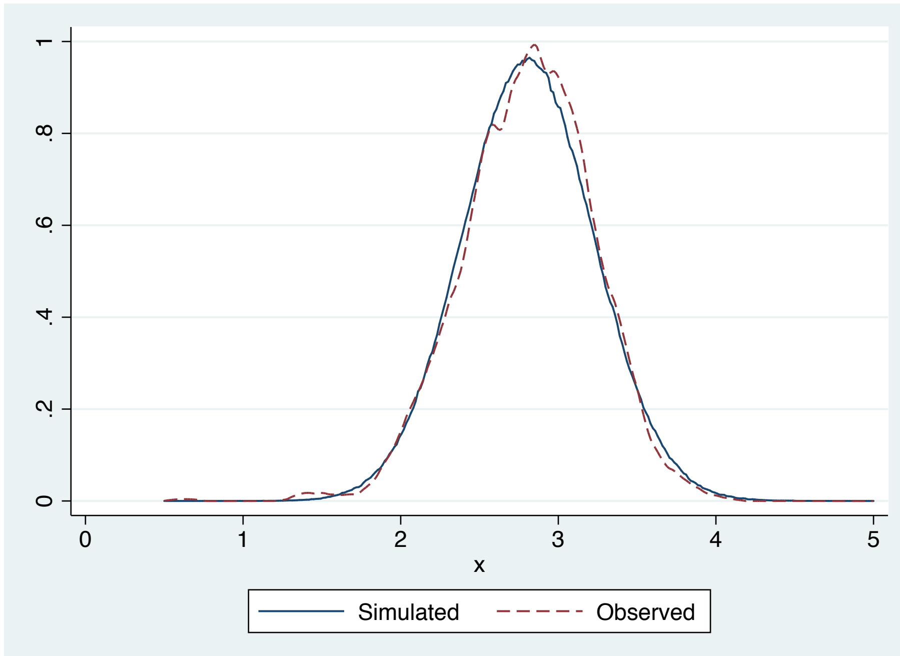
	Cognitive		**Mechanical Aux**		**Socio-emotional**	
	Estimate	SE	Estimate	SE	Estimate	SE
μ_1	-0.57	0.29	-0.39	0.12	1.05	0.11
μ_2	0.39	0.11	0.37	0.05	-0.53	0.07
$1/\sigma_1^2$	2.42	0.75	4.33	0.92	6.39	1.92
$1/\sigma_2^2$	4.26	1.14	12.54	2.77	4.15	1.26
p	0.44	0.19	0.50	0.10	0.34	0.05
1-p	0.56	0.19	0.50	0.10	0.66	0.05

Note: This table presents estimates from the Model. Since the model is estimated using Bayesian methods, they represent the mean estimates over 1,000 iterations after discarting the first 30,000.The computation of standard errors is explained in appendix 3. Mechanical Aux. presents the results from the auxiliar component of the factor, θ_2, that is independent from cognitive ability. Where $\theta_m = \alpha_1\theta_c + \theta_2$ with $\alpha_1 = 0.42$

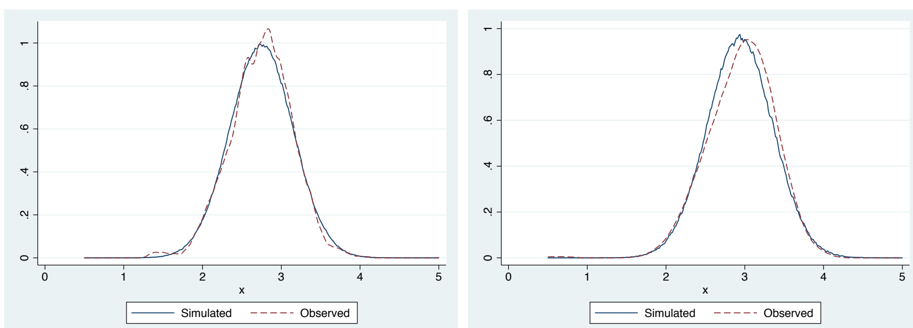
a better job predicting log wages than a two-factor model that does not include the mechanical factor. Both models predict well college attendance decisions.

Figure 4.16 compares the actual distribution of log wages with the distribution of the simulated log wages for the whole sample (panel a) and by schooling level, in panels b and c. The two distributions are very similar although the mean wage for individuals that attended college is lower than the observed mean. Table 4.15 presents a formal goodness of fit test for log wages wages. The chi-squared test cannot reject the null hypothesis that the simulated distribution of hourly wages is statistically equivalent to the actual distribution observed in the data.

Figure 4.16: Simulated versus Observed Wages



(a) Overall



(b) No college

(c) College

Note: The dashed line depicts the actual distribution of log hourly wage in the data while the solid line is computed after simulating a sample of over 1'000.000 individuals using the structure and estimates of the model.

Moreover, the three factor model used is superior than an alternative two factor model that does not take into account mechanical ability. In fact, the two factor model cannot succesfully reproduce the distribution of log hourly wages. Table 4.15 presents the results of the chi-squared goodness of fit test on the simulated distribution of hourly wages that corresponds to a model with three and two factors (only cognitive and socio-emotional). The null hypothesis for the model of two factors is rejected[12].

Table 4.15: Goodness of Fit: Wage Distribution

	3 factors	**2 factors**
χ^2	46.61	272.46
p-value	0.19	0.00
Critical at 90%	50.66	50.66
Critical at 95%	54.57	54.57

Note: The table presents a Chi-squared test computed using equiprobable bins. Ho:Model=Data

Finally, in Table 4.16 we compare the performace of our model with a model of two factors in predicting college attendace. In both cases the tests cannot reject the null hypothesis which implies that the two models present a good fit with the data.

Table 4.16: Goodness of Fit: Schooling

	3 factors	**2 factors**
χ^2	0.40	0.02
p-value	0.53	0.87
Critical at 90%	2.71	2.71
Critical at 95%	3.84	3.84

Note: The table presents a Chi-squared test. Ho:Model=Data.

[12] It is useful to point out that Heckman et al. (2006) find similar results when computing the Chi-squared test on the sample of 4-year college graduates.

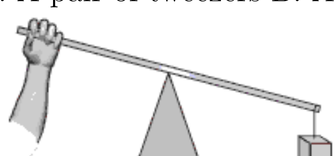
Chapter A: Appendix

A.1 Sample Questions

The set of questions was extracted from: http://www.education.com/reference/article/mechanical-comprehension-quiz/

A.1.1 Mechanical Comprehension Section

1. The diagram shows a class 1 lever. Which of the following is the same kind of lever? A. A pair of tweezers B. A pair of scissors C. A wheelbarrow D. A pair

   of tongs

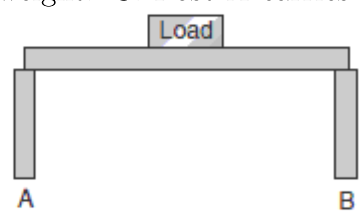
2. The diagram shows a class 2 lever. Which of the following is the same kind of lever? A. A seesaw B. A pair of scissors C. The human forearm D. A wheelbarrow

3. When a mass of air expands, which of the following is most likely to happen? A. The air warms up. B. The air cools down. C. The air stays at the same temperature. D. The air contracts.

4. The diagram shows a class 3 lever. Which of the following is the same kind of lever? A. A pair of tweezers B. A wheelbarrow C. A seesaw D. A wedge

5. Which of the following would feel hottest to the touch if one end were placed in a pot of boiling water? A. A wooden spoon B. A metal fork C. A plastic knife D. A plastic cup

6. In the diagram, what can you tell about the load on posts A and B? A. Post B carries more weight. B. Post A carries more weight. C. Post A carries no weight. D. The load is equal on posts A and B.



7. Water is flowing through this pipe. Which statement is true? A. Water is moving faster at point A than at point B. Water pressure is equal at points A and B. C. Water pressure is greater at point A than at point B. D. Water pressure is greater at point B than at point A.

8. What is the advantage of using triangle shapes in constructing a bridge? A. Triangles are sturdier than other shapes. B. Triangles are very flexible. C. Triangles are inexpensive to manufacture. D. Triangles are attractive to look at.

9. Shifting to a smaller gear on a mountain bike will have an effect on the speed of travel. The smaller sized gear will make pedaling easier but it will also a. increase the speed of travel. b. decrease the speed of travel. c. have no effect

on the speed of travel. d. make the bicyclist work harder.

10. Which of the following examples does not make use of a wedge? a. Choosing a sand wedge to hit your golf ball b. Splitting firewood with a chisel and sledge hammer c. Chopping wood with an axe d. Using a lever to lift a load

11. A block and tackle refers to a device which is used to a. put under the wheel of a vehicle to prevent it from rolling backward. b. prevent fish from escaping the hook. c. leverage a stationary object. d. hoist an object into the air by means of rope and pulleys.

12. Downshifting an auto or a truck causes a. a decrease in speed and an increase in torque. b. an increase in speed and a decrease in torque. c. no change in speed and no change in torque. d. None of the above

13. Shifting to a higher gear in a car or truck causes a. a decrease in torque and an increase in speed. b. an increase in torque and a decrease in speed. c. an increase in both speed and torque. d. None of the above.

A.1.2 Automotive and Shop Information

1. A car uses too much oil when which of the following parts are worn? A. pistons B. piston rings C. main bearings D. connecting rods

2. What system of an automobile or truck determines the vehicle's cornering ability and ride stiffness? a. Steering system b. Braking system c. Electrical system d. Suspension system

3. The purpose of a transfer case is to a. make a vehicle ride more smoothly. b. make the steering more responsive to driver input. c. distribute power to front and rear wheels in a 4 x 4 vehicle. d. shorten the braking distance.

4. The reason a particular quarter inch nut may not fit a particular quarter inch bolt is because a. they may be of different thread classifications. b. a quarter inch bolt is incompatible with a quarter inch nut of the same size. c. the metal alloys from which the nut and bolt are made may cause the nut to seize.d. quarter-inch bolts require a nut of a slightly larger size to fit.

5. The kerf is a. a type of wood file. b. the angle of the blade on a circular saw. c. a slot or cut made by the blade of a saw as it cuts into the wood. d. a term of measurement used in vehicle wheel alignment.

6. It would be better to use thick viscosity motor oil in a. cold climates (makes vehicle startups easier). b. tropical climates (engine heat build-up). c. Eastern United States. d. four-wheel drive vehicles.

7. The part of the motor vehicle electric system which distributes the spark to the various combustion cylinders is the a. battery. b. rotor and distributor assembly. c. injection system. d. ignition coil.

8. A punch is used for a. hammering knots from wooden objects. b. marking metal or wooden objects to prepare for drilling or other activities and for driving small headed nails. c. filing the sharp edges of metal or wooden objects. d. drilling holes.

9. For a better grip on a stubborn fastener nut, it is better to use a. an adjustable wrench. b. an open-end wrench. c. a box-end wrench. d. a pipe wrench.

A.1.3 Electronics Information

1. Ohm's Law states that a. E = I x R. b. R = E x I. c. voltage is equal to the current multiplied by the resistance. d. Both a and c

2. The electrons revolve around the nucleus in a cumulative series of orbits which are called a. neutrons. b. subatomic particles. c. shells. d. circulating cores.

3. The part of the atom's shell that determines electrical properties is the __________ shell. a. insulator b. nucleic c. valence d. electronic

4. A semi-conductor is an element or substance which a. conducts electricity better than a conductor. b. is useful for certain conductive requirements necessary to some electrical technologies. c. completely inhibits the flow of electrons around the outer shell. d. insulates electrical current from contact with other materials.

5. When applied to electrical conductivity of household current, 60 hertz means that a. current flows in only one direction. b. current flows in two directions. c. current flows first in one direction and then another. d. 60 voltage cycles take place in one second.

6. The three necessary components of an electrical circuit are a. an electrical load, conductors, and a circuit for the electricity flow to follow. b. a switch, a

resistor, and a path to follow. c. a 60 hertz receptacle, a switch, and a power source. d. a closed circuit, a battery, and radio waves.

7. Doping is a term used in the semiconductor process when a. impurities are added into the crystal structure of silicon. b. hydrogen atoms are added to the crystal structure of silicon. c. impurities are removed from the crystal structure of silicon. d. semiconductors are used for medical purposes.

8. The property of electricity that pushes and moves it along a circuit is called a. alternating current. b. amperage. c. resistance. d. voltage.

Chapter B: Appendix

B.1 Identification of the Model

This section presents the identification of the empirical model utilized in chapter 2. the identification of the model used in chapter 3 follows the same rationale. I follow Carneiro et al. (2003). For notational simplicity, I keep the conditioning on X implicit and focus on the factors (latent abilities).

Let C_j denote the cognitive test scores

$$C_j = \lambda^c_{C_j}\theta_c + e_{C_j}$$

for $j = 1, ..., 6$

where θ_c is the cognitive factor, $\lambda^c_{C_j}$ is the loading of the cognitive factor in test j and e_{C_j} is the error term (uniquenesses).

I can compute

$$\begin{aligned}
COV(C_1, C_2) &= \lambda^c_{C_1}\lambda^c_{C_2}\sigma^2_{\theta_c} \\
COV(C_1, C_3) &= \lambda^c_{C_1}\lambda^c_{C_3}\sigma^2_{\theta_c} \\
COV(C_2, C_3) &= \lambda^c_{C_2}\lambda^c_{C_3}\sigma^2_{\theta_c}
\end{aligned}$$

Since I observe the left hand side, I can form

$$\frac{COV(C_1, C_2)}{COV(C_2, C_3)} = \frac{\lambda^c_{C_1}}{\lambda^c_{C_3}}$$
$$\frac{COV(C_1, C_2)}{COV(C_1, C_3)} = \frac{\lambda^c_{C_2}}{\lambda^c_{C_3}}$$

By normalizing $\lambda^c_{C_3} = 1$, I get $\lambda^c_{C_1}$ and $\lambda^c_{C_2}$. With this I can also get $\sigma^2_{\theta_c}$ and apply the same procedure for the rest of the tests C_4, C_5, C_6.

Finally, I can rewrite the system as:

$$\frac{C_j}{\lambda^c_{C_j}} = \theta_c + \frac{\varepsilon_{C_j}}{\lambda_{C_j}} = \theta_c + \varepsilon'_{C_j}$$

and I can apply Kotlarski's Theorem (Kotlarski, 1967) to identify

$$f_{\theta_c}(\cdot), f_{\varepsilon_{C_j}}(\cdot)$$

for $j = 1, ..., 6$

To implement the model I need to assume $\lambda_{C_j} = 1$ for some j. This assumption sets the scale of θ_c. In this case I set the scale of unobserved cognitive ability by normalizing to one the coefficient associated with θ_c in the equation for mathematics knowledge.

For the identification of the distribution of socio-emotional ability I use a similar argument. In particular, consider the two noncognitive test scores and the latent variable associated with the schooling model.

$$S_1 = \lambda^s_{S_1}\theta_s + e_{S_1}$$

$$S_2 = \lambda^s_{S_2}\theta_s + e_{S_2}$$

$$I = \lambda^c_D\theta_c + \lambda^m_D\theta_m + \lambda^s_D\theta_s + e$$

Given that $\theta_c \perp \theta_s$ and $\theta_m \perp \theta_s$, I can compute

$$\begin{aligned} COV(S_1, I) &= \lambda^s_{S_1}\lambda^s_D\sigma^2_{\theta_s} \\ COV(S_2, I) &= \lambda^c_{S_2}\lambda^s_D\sigma^2_{\theta_c} \end{aligned}$$

and

$$\frac{COV(S_1, I)}{COV(S_2, I)} = \frac{\lambda^s_{S_1}}{\lambda^s_{S_2}}$$

so the normalization $\lambda^s_{S_1} = 1$ ensures the identification of the loading $\lambda^s_{S_2}$. With $\lambda^s_{S_2}$ in hand, I secure the identification of the distribution of θ_s using Kotlarski's theorem. In this case I normalize the coefficient associated with θ_s in the equation for the Rosenberg Self-Esteem Scale.

Finally, for the mechanical measure M_k I have to consider that both θ_c and θ_m are present in the equations and they are not independent. In order to use the same chain logic applied to the identification of the other to factors I rewrite the system in terms of two independent factors. For this purpose I assume that

$$\theta_m = \alpha_1\theta_c + \alpha_2\theta_2$$

where

$$\theta_c \perp \theta_2$$

and both θ_c and θ_2 are distributed as a mixture of normals as follows:

$$\theta_{c,i} \sim \sum_{k=1}^{K} p_k N\left(\mu_c^k, \left(\sigma_c^k\right)^2\right)$$

$$\theta_{m,i} \sim \sum_{j=1}^{J} p_j N\left(\mu_m^j, \left(\sigma_m^j\right)^2\right)$$

Without loss of generality I assume $\alpha_2 = 1$ so I normalize the contribution of θ_c to θ_m. So the original model for the mechanical measure can be rewritten in terms of θ_c and θ_2 as follows:

$$\begin{aligned} M_k &= \lambda_{M_k}^c\theta_c + \lambda_{M_k}^m\theta_m + e_{M_k} \\ &= \lambda_{M_k}^c\theta_c + \lambda_{M_k}^m(\alpha_1\theta_c + \theta_2) + e_{M_k} \\ &= a_k\theta_c + \lambda_{M_k}^m\theta_2 + e_{M_k} \end{aligned}$$

for $k = 1, ..., 3$

I can compute

$$
\begin{aligned}
COV(C_1, M_1) &= \lambda^c_{C_1} a_1 \sigma^2_{\theta_c} \\
COV(C_1, M_2) &= \lambda^c_{C_1} a_2 \sigma^2_{\theta_c} \\
COV(C_1, M_3) &= \lambda^c_{C_1} a_3 \sigma^2_{\theta_c}
\end{aligned}
$$

to recover a_1, a_2 and a_3.

As for the other test scores, I normalize $\lambda^m_{M_3} = 1$. To apply Klotarski's Theorem I rewrite the system as:

$$
\begin{aligned}
\frac{M_1 - a_1\theta_c}{\lambda^m_{M_1}} &= \theta_2 + e^{'}_{M_1} \\
\frac{M_2 - a_2\theta_c}{\lambda^m_{M_2}} &= \theta_2 + e^{'}_{M_2} \\
M_3 - a_3\theta_c &= \theta_2 + e^{'}_{M_3}
\end{aligned}
$$

and I identify the the distribution of $f_{\theta_2}(\cdot)$, $f_{e_{M_k}}(\cdot)$ for $k = 1, 2, 3$

Finally, to recover all the parameters associated with θ_m I need to get α_1 so one extra assumption is needed since I have three equations and four unknowns in the following system:

$$a_1 = \lambda^c_{M_1} + \lambda^m_{M_1}\alpha_1$$

$$a_2 = \lambda^c_{M_2} + \lambda^m_{M_2}\alpha_1$$

$$a_3 = \lambda^c_{M_3} + \alpha_1$$

I assume that $\lambda^c_{M_1} = 0$, the implication of the assumption is that the cognitive factor θ_c affects the score only through its effect on the mechanical factor θ_m[1]

In the implementation of the model I normalize to one the coefficient associated with θ_m in the equation for mechanical comprehension.

B.2 Standard Errors of the Estimates

In order to justify the computation of standard errors presented in this paper it is necessary to introduce some Bayesian concepts and the corresponding notation.

Let θ be the parameter of interest in this case $\theta = (\alpha, \beta, \lambda)$, $f(\theta)$ the density of θ, called the prior distribution. $Y = \{y_{1,..,}y_N\}$ is the sample of N independent observations, where $f(y_n|\theta)$ is the probability of outcome y_n, and $f(Y)$ the marginal distribution of the data (marginal over θ). The posterior distribution is denoted by $f(\theta|Y)$ and the probability of observing the sample outcomes Y is the likelihood function of the observed choices $L(Y|\theta) = \prod_{i=1}^{N} f(y_n|\theta)$.

[1] In the implementation of the model M_1 is the score associated with the automotive and shop information section. I selected this test because it has the loIst loading on the cognitive factor in the premilinary factor analysis (see 4.2) The current results do not depend on this assumptions, results are qualitatively similar if I select any section on the technical composites of the ASVAB (mechanical comprehension or electronics information). Results are available upon request.

In this context $f(Y) = \int L(Y|\theta)f(\theta)d\theta$ and using the Bayes' rule the following equality is true and serves to compute the desired posterior distribution of θ.

$$f(\theta|Y)f(Y) = L(Y|\theta)f(\theta)$$

$$f(\theta|Y) = \frac{L(Y|\theta)f(\theta)}{f(Y)}$$

$$f(\theta|Y) \propto L(Y|\theta)f(\theta)$$

Finally, the mean of the posterior distribution is

$$\bar{\theta} = \int \theta f(\theta|Y)d\theta \tag{B.1}$$

The use of Bayesian methods in this paper is merely computational; in consequence, the interest is primarily on the mean of the posterior distribution $\bar{\theta}$ which is vieId from a classical perspective, i.e., as an estimator that has the same asymptotic sampling distribution as the maximum likelihood estimator.[2] In this sense, the interest is to find the sampling distribution of the statistic$\bar{\theta}$in order to make inference about it.

The Bernstein-von Mises theorem, described by Train (2003) in three related statements establishes the properties of the sampling distribution of $\bar{\theta}$:

[2]From a bayesian perspective, the mean of the posterior distribution is the value that minimizes the posterior loss in the quadratic loss case. As stated in Train (2003) is the value that minimizes the expected cost of the researcher being wrong about the parameter, if the cost is quadratic in the size of the error.

1. $\sqrt{N}(\theta - \bar{\theta}) \rightarrow^{d} N(0, (-H)^{-1})$[3]

2.$\sqrt{N}(\bar{\theta} - \theta^{MLE}) \rightarrow^{p} 0$

3.$\sqrt{N}(\bar{\theta} - \theta^{*}) \rightarrow^{d} N(0, (-H)^{-1})$

In this context, the variance of the posterior is the asymptotic variance of the estimates. From 1 I have that the asymptotic variance of the posterior distribution is $(-H)^{-1}/N$ which by 3 is the asymptotic sampling variance of the estimator $\bar{\theta}$. So, estimation can be performed by using the moments of the posterior, as in this paper, where the mean of the posterior provides a point estimate and the standard deviation of the posterior provides the standard errors.

In the paper, I use MCMC as a method to obtain draws from the posterior distribution. Starting with a vector of initial parameters drawn from the transition kernel, I use Gibbs Sampling as the algorithm to create a Markov Chain such that, as size of the sequence increases ($n \rightarrow \infty$), the limiting distribution is the posterior. After convergence is achieved and a burning period of 60,000, I make 1,000 draws from the posterior distribution of the parameters to compute the mean (the simulated approximation of the mean $\bar{\theta}$ that I call $\breve{\theta}$) and standard errors (provided by the sd of the posterior which is simulated by taking the the standard deviation of the R draws) reported in the text.

$$\breve{\theta} = \frac{\sum_{r=1}^{R} \theta^{r}}{R}$$

[3]With $-H$ being the information matrix (the negative)

$$SE_{\breve{\theta}} = \sqrt{\frac{\sum_{r=1}^{R}(\theta^r - \bar{\theta})^2}{R}}$$

According to Gelman and Shirley (2011) when simulation-based inference is for functions of the parameters $g(\theta)$. "Such inference will typically be constructed using a collection of 1000 (say) simulations of the parameter vector, perhaps summarized by a mean and standard deviation, or maybe a 95% interval using the empirical distribution of the simulations that have been saved. Even if these summaries could be computed analytically, I would in general still want simulations because these allow us directly to obtain inferences for any posterior or predictive summary".

Bibliography

Acemoglu, D. and D. H. Autor (2011). kills, tasks and technologies: Implications for employment and earnings. In O. Ashenfelter and D. E. Card (Eds.), *Handbook of Labor Economics*, Volume 4, Chapter 12, pp. 1043?1171. Amsterdam: Elsevier.

Altonji, J. (2010). Multiple skills, multiple types of education, and the labor market: A research agenda. note AEA: Ten Years and Beyond: Economists Answer NSF's Call for Long-Term Research Agendas.

Autor, D. and H. Michael (2013). Putting tasks to the test: Human capital, job tasks, and wages. *Journal of Labor Economics 31*(2), S59–S56.

Autor, D. H. and D. Dorn (2009). The growth of low skill service jobs and the polarization of the u.s. labor market. Working Paper 15150, National Bureau of Economic Research.

Autor, D. H., L. F. Katz, and M. S. Kearney (2006, May). The polarization of the U.S. labor market. *American Economic Review 96*(2), 189–194.

Autor, D. H., F. Levy, and R. J. Murnane (2003, November). The skill content of recent technological change: An empirical exploration. *Quarterly Journal of Economics 118*(4), 1279?1333.

Bishop, J. (1988). Occupational competency as a predictor of labor market performance. Cornell University, School of Industrial and Labor Relations.

Bishop, J. H. (1991). Achievement, test scores, and relative wages. In M. H. Kosters (Ed.), *Workers and their wages: Changing patterns in the United States*, Number 520 in AEI Studies, pp. 146–186. Washington, D.C.: AEI Press.

Blauvelt, G. R. (2006). *MachineShop: A Design Environment for Supporting Children's Construction of Mechanical Reasoning and Spatial Cognition*. Ph. D. thesis, University of Colorado.

Boehm, M. J. (2013, September). The wage effects of job polarization: Evidence from the allocation of talents. http://personal.lse.ac.uk/boehm/.

Borghans, L., A. L. Duckworth, J. J. Heckman, and B. ter Weel (2008, Fall). The economics and psychology of personality traits. *Journal of Human Resources 43*(4), 972–1059.

Borghans, L., B. ter Weel, and B. A. Weinberg (2007, January). Interpersonal styles and labor market outcomes. Working Paper 12846, NBER.

Bowles, S. and H. Gintis (1976). *Schooling in Capitalist America: Educational Reform and the Contradictions of Economic Life*. New York: Basic Books.

Bowles, S., H. Gintis, and M. Osborne (2001, May). Incentive-enhancing preferences: Personality, behavior, and earnings. *American Economic Review 91*(2), 155–158.

Cameron, S. V. and J. J. Heckman (1998, April). Life cycle schooling and dynamic selection bias: Models and evidence for five cohorts of American males. *Journal of Political Economy 106*(2), 262–333.

Cameron, S. V. and J. J. Heckman (2001, June). The dynamics of educational attainment for black, hispanic, and white males. *Journal of Political Economy 109*(3), 455–99.

Carneiro, P., K. Hansen, and J. J. Heckman (2003, May). Estimating distributions of treatment effects with an application to the returns to schooling and measurement of the effects of uncertainty on college choice. *International Economic Review 44*(2), 361–422.

Carneiro, P. and J. J. Heckman (2003). Human capital policy. In J. J. Heckman, A. B. Krueger, and B. M. Friedman (Eds.), *Inequality in America: What Role for Human Capital Policies?*, pp. 77–239. Cambridge, MA: MIT Press.

Carpenter, P. A. and M. A. Just (1989, July). Comprehension processes in mechanical reasoning. Technical report, Carnegie Mellon University.

Cawley, J., J. J. Heckman, and E. J. Vytlacil (2001, September). Three observations on wages and measured cognitive ability. *Labour Economics 8*(4), 419–442.

Conti, G., J. J. Heckman, and S. Urzua (2010, May). The education-health gradient. *American Economic Review: Papers and Proceedings 100*(2), 1–5.

Cox, J. (1928). *Mechanical Aptitude: Its Existence, Nature, and Measurement.* Methuen and Company Ltd.

Cunha, F. and J. J. Heckman (2007, May). The technology of skill formation. *American Economic Review 97*(2), 31–47.

Cunha, F., J. J. Heckman, L. J. Lochner, and D. V. Masterov (2006). Interpreting the evidence on life cycle skill formation. In E. A. Hanushek and F. Welch (Eds.), *Handbook of the Economics of Education*, Chapter 12, pp. 697–812. Amsterdam: North-Holland.

Cunha, F., J. J. Heckman, and S. M. Schennach (2010, May). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica 78*(3), 883–931.

Dickens, W. T. (2008). *Cognitive Ability.* The New Palgrave Dictionary of Economics. Steve Durlauf ed.

Diekhoff, G. (1992). *Statistics for the social and behavioral sciences: Univariate, bivariate, multivariate.* Wm. C. Brown Publishers (Dubuque, IA).

Duckworth, A. and S. Urzua (2009). Determinants of success in early adulthood: Comparing the effects of intelligence and big five personality traits. Unpublished manuscript, University of Pennsylvania, Department of Psychology.

Duckworth, A. L., C. Peterson, M. D. Matthews, and D. R. Kelly (2007, June). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology 92*(6), 1087–1101.

Duncan, G. J. and R. Dunifon (1998). "soft-skills" and long-run labor market success. In S. W. Polachek (Ed.), *Research in labor economics*, Volume 17, pp. 123–149. Stamford, CT: JAI Press.

Edwards, R. C. (1976, Winter). Individual traits and organizational incentives: What makes a 'good' worker? *Journal of Human Resources 11*(1), 51–68.

Ellwood, D. T. and T. J. Kane (2000). Who is getting a college education? Family background and the growing gaps in enrollment. In S. Danziger and J. Waldfogel (Eds.), *Securing the Future: Investing in Children from Birth to College*, pp. 283–324. New York: Russell Sage Foundation.

Farkas, G. (2003, August). Cognitive skills and noncognitive traits and behaviors in stratification processes. *Annual Review of Sociology 29*, 541–562.

Ferguson, E., J. J. Heckman, and P. Corr (2011). Editorial: Personality and economics: Overview and proposed framework. *Personality and Individual Differences 51*(Special Issue on Personality and Economics), 201–209. E. Ferguson, J.J. Heckman, and P. Corr, editors.

Ferguson, T. S. (1983). Bayesian density estimation by mixtures of normal distributions. In H. Chernoff, M. Rizvi, J. Rustagi, and D. Siegmund (Eds.), *Recent Advances in Statistics: Papers in Honor of Herman Chernoff on his Sixtieth Birthday*, pp. 287–302. New York: Academic Press.

Firpo, S., M. F. Nicole, and T. Lemieux (2011, February). Occupational tasks and changes in the wage structure. IZA Discussion Paper No. 5542.

Gelman, A. and K. Shirley (2011). Inference from simulations and monitoring convergence. In G. J. S. Brooks, A. Gelman and X. L. Meng (Eds.), *Handbook of Markov Chain Monte Carlo*. CRC Press.

Gibbons, R., L. Katz, T. Lemieux, and D. Parent (2005). Comparative advantage, learning, and sectoral wage determination. *Journal of Labor Economics 23*(4), 681–724.

Hansen, K. T., J. J. Heckman, and K. J. Mullen (2004, July–August). The effect of schooling and ability on achievement test scores. *Journal of Econometrics 121*(1-2), 39–98.

Hartog, J. (2001, December). On human capital and individual capabilities. *Review of Income and Wealth 47*(4), 515–540.

Hartog, J. M. v. P. and J. V. D. Sluis (2010, Winter). If you are so smart, why aren't you an entrepreneur? returns to cognitive and social ability: Entrepreneurs versus employees. *Journal of Economics and Management Strategy 19*(4), 947–989.

Hause, J. C. (1972). Earnings profile, ability and schooling. *Journal of Political Economy 80*(3), S108?S138.

Heckman, J. J. (1995). Notes on schooling, earnings and ability. Unpublished manuscript, University of Chicago, Department of Economics.

Heckman, J. J. and T. Kautz (2013). Fostering and measuring skills: Interventions that improve character and cognition. Working Paper 19656, National Bureau of Economic Research. Download from http://www.nber.org/papers/w19656.

Heckman, J. J., J. Stixrud, and S. Urzua (2006, July). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics 24*(3), 411–482.

Hegarty, M. (1992). Mental animation: Inferring motion from static displays of mechanical systems. *Journal of Experimental Psychology 18*(5), 1084?1102.

Hegarty, M., M. A. Just, and I. R. Morrison (1988). Mental models of mechanical systems: Individual differences in qualitative and quantitative reasoning. *Cognitive Psychology 20*(2), 191–236.

Heiser, J. and B. Tversky (2002). Diagrams and depictions in acquiring complex systems. In W. D. Gray and C. Schunn (Eds.), *Proceedings of the 24th Annual Meeting of the Cognitive Science Society.*

Herrnstein, R. J. and C. A. Murray (1994). *The Bell Curve: Intelligence and Class Structure in American Life.* New York: Free Press.

Ingram, B. F. and G. R. Neumann (2006, February). The returns to skill. *Labour Economics 13*(1), 35–59.

Jencks, C. (1979). *Who Gets Ahead? The Determinants of Economic Success in America.* New York: Basic Books.

Keane, M. P. and K. I. Wolpin (1997, June). The career decisions of young men. *Journal of Political Economy 105*(3), 473–522.

Kotlarski, I. I. (1967). On characterizing the gamma and normal distribution. *Pacific Journal of Mathematics 20*, 69–76.

Lazear, E. (2003). Firm-specific human capital: A skill-weights approach. NBER Working Papers 9679.

Lee, D. (2005, February). An estimable dynamic general equilibrium model of work, schooling, and occupational choice. *International Economic Review 46*(1), 1–34.

Menaghan, E. G. (1990). The impact of occupational and economic pressures on young mothers' self-esteem: Evidence from the nlsy. Presented: Annual Meetings of the Society for the Sociological Study of Social Problems, Washington, D.C., August 9.

Miller, A. (1980). Work, jobs and occupations: A critical review of the dictionary of occupational tittles. Technical report, National Academy Press.

Murnane, R. J., J. B. Willett, and F. Levy (1995, May). The growing importance of cognitive skills in wage determination. *Review of Economics and Statistics 77*(2), 251–266.

NCES (2013). The condition of education 2013 (nces 2013-037).

Neal, D. A. and W. R. Johnson (1996, October). The role of premarket factors in black-white wage differences. *Journal of Political Economy 104*(5), 869–895.

O'Neill, J. (1990). The role of human capital in earnings differences between black and white men. *The Journal of Economic Perspectives 4*(4), 25–45.

P., B. M. and B. S. Blu (2010). Two sides of the same coin u.s. residual inequality and the gender gap. *Journal of Human Resources 45*(1), 197–242.

Paterson, D. G., R. Elliot, L. D. Anderson, H. A. Toops, and E. Heidbreder (1930). Minnesota mechanical ability tests: The report of a research investigation subsidized by the committee on human migrations of the national research council. Technical report, Department of Psychology of the University of Minnesota.

Poletaev, M. and C. Robinson (2008, July). Human capital specificity: Evidence from the dictionary of occupational titles and displaced worker surveys 1984-2000. *Journal of Labor Economics 26*(3), 387–420.

Rosenberg, M. (1965). *Society and the Adolescent Self-Image*. Princeton, NJ: Princeton University Press.

Rotter, J. B. (1966). *Generalized Expectancies for Internal versus External Control of Reinforcement*. Washington DC: American Psychological Association.

Roy, A. (1951, June). Some thoughts on the distribution of earnings. *Oxford Economic Papers 3*(2), 135–146.

Rubinstein, Y. and Y. Weiss (2006). Post schooling wage growth: Investment, search and learning. In E. Hanushek and F. Welch (Eds.), *Handbook of the Economics of Education*, Volume 1 of *Handbooks in Economics*, Chapter 1, pp. 1–67. Amsterdam: North-Holland.

Sarzosa, M. and S. Urzua (2013). Bullying and cyberbullying in teenagers: The role of cognitive and non-cognitive skills. University of Maryland.

Sheskin, D. (2004). *Handbook of Parametric and Nonparametric Statistical Procedures* (Third Edition ed.). Chapman and Hall/CRC Press.

Spearman, C. (1904). "general intelligence," objectively determined and measured. *American Journal of Psychology 15*, 201–293.

Spearman, C. (1923). *The nature of intelligence and the principles of cognition.* London: Macmillan.

Stenquist, J. L. (1923). *Measurements of Mechanical Ability.* Columbia University Teachers College.

Strocchia-Rivera, L. (1988). *Self-Esteem and Educational Aspirations as Antecedents of Adolescent Unmarried Motherhood.* Ph. D. thesis, The University of Texas at Austin.

Sullivan, P. (2010, February). A dynamic analysis of educational attainment occupational choices and job search. *International Economic Review 51*(1), 289–317.

Tambunlerchai, S. (2011). *Choosing Entrepreneurship: The Roles of Cognitive and Nonocognitive Abilities on Self Employment Decisions and Outcomes.* Ph. d. thesis, Department of Economics, University of Chicago, Chicago, IL.

Train, K. E. (2003). *Discrete Choice Methods with Simulation.* National Economic Research Associates, Inc: Cambridge University Press.

Urzua, S. (2008, Fall). Racial labor market gaps: The role of abilities and schooling choices. *Journal of Human Resources 43*(4), 919–971.

Willis, R. J. (1986). Wage determinants: A survey and reinterpretation of human capital earnings functions. In O. Ashenfelter and R. Layard (Eds.), *Handbook of Labor Economics*, Volume 1, pp. 525–602. New York: North-Holland.

Willis, R. J. and S. Rosen (1979, October). Education and self-selection. *Journal of Political Economy 87*(5, Part 2), S7–S36.

Wise, L., J. Welsh, F. Grafton, F. P., J. Earles, L. Sawin, and D. R. Divgi (1992). Sensitivity and fairness of the armed services vocational aptitude battery (asvab) technical composites. Technical report, Seaside, CA: Defense Manpower Data Center.

Wolfe, R. N. and S. D. Johnson (1995, April). Personality as a predictor of college performance. *Educational And Psychological Measurement 55*(2), 177–185.

Yamaguchi, S. (2012, January). Tasks and heterogeneous human capital. *Journal of Labor Economics, pp. 1-53 30*(1), 1–53.