Beyond qualifications: labor market returns to cognitive and socio-emotional skills in urban Colombia[1]

Pablo Acosta (World Bank)[2]
Noël Muller (World Bank)
Miguel Sarzosa (University of Maryland)

This version: September 13, 2014

Abstract

This paper estimates the returns to skills on labor outcomes for the working-age population living in metropolitan areas of Colombia. The analysis is based on an original survey collected in 2012 by the World Bank which provides measures of cognitive skills (a scaled reading proficiency assessment) and socio emotional-skills (personality traits and behaviors). The paper uses conventional statistical approaches and an advanced structural Maximum Likelihood Estimation method to estimate the effect of latent skills and traits to fully capture their influence on labor market outcomes. Findings indicate a stronger role for reading proficiency than for socio-emotional skills as a predictor of favorable labor outcomes (labor earnings, holding a formal job, being a white collar worker). Socio-emotional skills appear to play a stronger role on labor participation decisions, but weaker on wages. Both reading proficiency, to a great extent, and socio-emotional skills, to a lesser extent, are associated with having attended tertiary education. Stronger effects are observed for subsamples such as women and younger individuals.

JEL Codes: J24, J31, I24.

Key words: Colombia, returns to skills, cognitive skills, socio-emotional skills, personality traits, behaviors, schooling.

[1] The authors are strongly indebted to Wendy Cunningham, Sergio Urzúa, Mónica Parra Torrado, Juan D. Barón, Ana Maria Oviedo, Natalia Millian, Alexandria Valerio, Maria Laura Sanchez Puerta, Sebastian Monroy, Tania Rajadel for their comments, suggestions, and frequent interactions with the authors while developing this paper.
[2] Contact author: pacosta@worldbank.org

1. Introduction

A central priority for policy makers is to increase the productivity of their country's economy and provide economic opportunities to their population by fostering the employability of the labor force. A better use and investment on skills have great implications regarding to labor productivity, innovation and growth (Banerji et al. 2010). Beyond that, providing vulnerable workers and the young with the type of skills that would improve their employability, cope with the economic transformations occurring in a global economy, and ultimately help them to find a quality job with sufficient income, are central policy issues from a human development perspective.

The main objective of this paper is to investigate (i) the current levels and distribution of skills in the Colombian working-age population, (ii) which skills (or skill sets) are most important for labor market and school success, and (iii) through which channels they impact on outcomes. It complements empirical evidence from high-income countries (United States, Western European countries) and the rare cases for developing countries on the impact of cognitive skills, socio-emotional and technical skills on schooling and labor market outcomes. These skills encompass respectively intelligence, behaviors and attitudes, and skills specifically developed to perform a range of tasks. Deeper knowledge on which abilities to foster, at what age, and through which interventions, would have major implications for the design of efficient education policies and active labor market programs such as trainings.

There is substantial evidence that employers across the globe value more a certain set of skills, not only qualifications. Recent studies using employers' surveys for Latin America confirm that, while technical skills are valued, also are socio-emotional skills and higher-order cognitive skills (Bassi et al. 2012). However, these employers tend to lament that potential recruits lack an adequate set of skills (of which work ethics and teamwork stands out), especially for the young. A strong mismatch seems to exist between the skills required by employers and those provided to students (Aedo and Walker 2012, Bassi et al. 2012, Mourshed et al. 2012).

At least for the Latin American case, among other explanations, this disconnection could root in the low quality of education in the region. Recent evidence shows that while between 1990 and 2010 the proportion of the labor force in the region with at least secondary education increased from 40 to 60 percent, returns to secondary education completion fell throughout the last two decades in the vast majority of countries, while the 2000s saw a reversal in the increase in the returns to tertiary education experienced in the 1990s (Gasparini et al. 2011). While several elements could explain this pattern[3], a degradation of higher education is considered as one of the main factor at stake (Bassi and Urzúa 2010, Gasparini et al. 2011, Bassi et al. 2012, Castro

[3] Potential factors are: over-supply of workers with higher educational attainment, reduction in the demand for labor requiring higher educational level, modified institutional policies like minimum wage.

and Yamada 2012, Levy and Schady 2013). Looking at the current distribution of skills and traits rather than qualifications would bring important insights to this debate.

Colombia is not an exception in this regional picture. It is a country with one of the highest income inequality in the already most unequal region in the world (Lustig et al. 2013). The sources of labor income inequality are multifold (incentives structures, labor market barriers, etc.) but skills might be one central. The wage premium declined for primary and secondary education between 1997 and 2008. Returns to university education fell during between 1997 and 2003 but increased in 2003-2008 (Aedo and Walker 2012). Despite decent economic growth in the last decade, the unemployment rate is also the highest in the region, especially for the youth. While it might paradoxically be partly a good sign (some individual can afford to be unemployed, dedicate more time to look for better job that match better their skills), it also highlights that jobs are not filled, despite employers' demand.

This paper contributes to the literature by using a unique dataset revealing information on cognitive skills, socio-emotional skills, skills used at work, and a rich background information of the Colombia working-age population living in urban areas. This is the first of the kind in Colombia, and one of the fewest worldwide for developing countries. This data and analysis can thus go beyond most previous studies that use education as proxy for human capital.

2. Concepts and definitions

One outstanding fact from the literature addressing skills, be it economic or psychology, is the plethora of definitions and taxonomies surrounding this concept. Economists have long considered educational levels as the main aspect capturing skills. Later on, studies have recognized that competencies and abilities are part of a broader scope and are also influenced by extra-school factors such as family background, work, extra-curricular activities and environment. Nowadays, the term "skills" is used broadly to include "competencies, attitudes, beliefs and behaviors that are malleable (modifiable) across an individual's development and can be learned and improved through specific programs and policies" (Guerra et al., 2014). Nonetheless, a variety of denominations have been used to designate these aspects.

Cognitive skills are generally defined as intelligence or mental abilities. Two levels are generally distinguished and referred as lower- and higher- orders of intelligence. Lower-order cognitive skills are basic skills such as literacy and numeracy. It relates to *crystallized intelligence* (knowledge and developed skills)[4]. Higher-order cognitive skills can be defined as the ability to understand complex ideas, adapt to the environment, learn from experience, engage in various forms of reasoning and overcome obstacles through thinking (Neisser et al. 1996). It relates to *fluid intelligence* (ability to solve novel problems) (Horn 1970, Cattell 1987). Cognitive skills are

[4] The latter is more malleable than the former and it is affected by formal schooling as well as for other stimuli that enhance mental capacity (Díaz et al. 2012).

usually measured by Intelligence Quotient (IQ) tests, other tests of intelligence and standardized test scores (reading, math and science).

Social-emotional skills refer to a distinct set of skills that enable individuals to navigate interpersonal and social situations effectively (Guerra 2014et al, 2014). Economists commonly considered behavioral characteristics and personality traits under the umbrella of "non-cognitive skills" or leave the distinction unexplained. Socio-emotional skills understood here as behaviors and attitudes (e.g. commitment, discipline, ability to work in a team and determination), and personality traits (such as self-confidence, sociability, emotional stability, among others[5]). Personality traits are relatively consensually seen as broad facets defining an individual and are relatively stable over time (Borghans et al. 2008a, Almlund et al. 2011). They influence socio-emotional skills as reactions. A widely used method to capture an individual's facets is to perform a factor analysis of the Five Factor model – or Big Five model - based on Goldberg's questionnaire (Goldberg 1993). Each of the five personality factors - openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism (opposite of emotional stability) - summarizes a large number of distinct, more specific personality traits, behaviors, and beliefs[6]. Another popular attribute in the literature is Grit, a narrower trait capturing one's inclination and motivation to achieve long term goals. The perseverance of effort and consistency of interest is characteristic of high-achieving individuals (Duckworth et al. 2007).

Finally, technical skills, also called professional skills or vocational abilities, can be defined as those abilities that are associated with the specific knowledge to carry out one's occupation. Measures of technical skills derive typically from an observed assessment of a person performs a task and the related skills. Alternative measurements can be drawn from test batteries related to mechanical, psychomotor abilities and manual dexterity (Prada, 2014). The psychology literature defines technical skills as a sub-set of cognitive skills (Almlund et al. 2011).

[5] The economic literature refers to socio-emotional skills using the terms "behavioral skills", "life skills", "non-cognitive skills" and "soft skills". However, a distinction has to be made between them. Non-cognitive skills refers to a broad range of behaviors, abilities and traits that are not induced by intelligence or achievement; soft and life skills usually include more technical skills such as language fluency and computer literacy 2014(Guerra et al, 2014). It is also worth noting that psychologists argue that many of the abilities and traits that economists intend to capture by the term "non-cognitive skills" are a result of cognition (Borghans et al. 2008a).

[6] The predominance of the Big-Five traits in the literature masks the diversity of behaviors and personality traits characterized as non-cognitive skills. No less than one hundred forty different socio-emotional skills have been found in the literature (Cunnigham and Villaseñor 2014). In order to propose a more practical taxonomy, 2014Guerra et al (2014) have put together a more comprehensive set of socio-emotional skills under eight facets that are known to be linked to labor outcomes of individuals: the PRACTICE skills.

3. Literature review

a. The role of skills and traits on labor earnings

Cognitive abilities have long been seen as the dominant factor determining labor earnings in the United States. The primary approach to assess such relationship is based on longitudinal data and consists in linking cognitive test scores of high school students (like mathematics) with their wages once they have graduated from high school. In a large number of economic studies, higher measures of cognitive skills were associated with higher wages and the capacity to deal with complex information processing in a professional environment (Herrnstein and Murray 1994, Murnane et al. 1995, Gottfredson 1997, Mulligan 1999, Murnane et al. 2000, Altonji and Pierret 2001, Cawley et al. 2001, Lazear 2003, Hanushek and Woessmann 2008). The net impact of measured cognitive abilities on earnings has been found high when taking into account differences in the quantity of schooling, workers' experience, and other factors that might also influence earnings (Hanushek and Woessmann 2008). Similar results were found in other high-income countries like the United Kingdom (McIntosh and Vignoles 2001) and Canada (Finnie and Meng 2001).

More recently, a burgeoning literature reported socio-emotional abilities and personality traits to be of equal importance or more in the determination of labor earnings of US workers.[7] Among the so-called Big-Five traits used in most empirical studies, conscientiousness is the most associated with job performance (Nyhus and Pons 2005, Almlund et al. 2011). Controlling for measures of cognition, empirical studies found that conscientiousness and traits related to emotional stability (locus of control[8] and self-esteem) play an essential role in determining job performance and wage (Bowles et al. 2001b, Judge and Hurst 2007; Drago 2011). Segal (2012) finds that misbehaviors of eight-grade students in 1988 – as measured by professor assessments - have been also found to be associated with lower earnings later in 1999 in the United States. Likewise, Kuhn and Weinberger (2005) suggest that males who occupied leadership positions in high school earn between 4 percent and 33 percent higher wages as adults.

Studies on high-school second chance programs in the United States have provided ideal natural experiments and their evaluation confirm the crucial role of socio-emotional skills. High-school dropout students who benefit from a second-chance program called the GED[9] exhibit higher levels of low-order cognitive skills than other high-school dropouts. However, it results that they

[7] A non-exhaustive list of sources claiming this results in the United States and Western Europe is: Bowles et al. (2001a), Nyhus and Pons (2005), Osborne-Groves (2005), Heckman et al. (2006), Mueller and Plug (2006), Carneiro et al. (2007), Borghans et al. (2008b), Kniesner and ter Weel (2008), Heineck and Anger (2010), Lindqvist and Vestman (2011), Segal (2012).

[8] The locus of control is defined as "the extent to which individuals believe they have control over their lives, i.e., self-motivation and self-determination (internal control) as opposed to the extent that the environment (i.e., chance, fate, luck) controls their lives (external control)" (Heckman et al. 2006).

[9] The General Educational Development (GED) is a certification program that aims to provide enrolled high school dropout students with equivalent qualifications than high school graduates.

have relative lower labor market outcomes later in life. In absolute terms, GED graduates have higher hourly wages but when controlling for ability, they earn less than other high-school dropouts. This unexplained factor is considered to originate from lower levels of socio-emotional skills that are also valued on the labor market. Being a GED graduate is a mixed signal that characterizes its recipients as smart but unreliable (Heckman and Rubinstein 2001).

Similar findings have been found for Western Europe countries. Men with lack of leadership abilities in Sweden – as measured by psychological assessment for the military assessment[10] - are more have lower annual earnings (Lindqvist and Vestman 2011). External locus of control (opposite of strong feeling of self-determination) has a strong negative impact on wage in Germany; however neuroticism has no robust effect on wages (Heineck and Anger 2010). Measures of behaviors and traits in childhood and adolescence also have important influence on adult earnings in West Germany and Great Britain (Carneiro et al. 2007, Borghans et al. 2008b).

There is much less evidence on the influence of cognitive and socio-emotional skills on labor earnings in Latin America, and developing countries in general. Cross-sectional data with measures of cognitive and socio-emotional skills, job performance and education of individuals in Argentina and Chile at the individual level have allowed drawing first evidence of that kind in the region (Bassi et al. 2012). The results suggest that self-efficacy[11] - a measure of high-level socio-emotional skills - is the ability that predominates in the association with higher wages in both countries. This association is strong at the post-secondary level but weak for workers with secondary traditional education. Other type of skills – low level socio-emotional skills, low- and high-order cognitive skills – do not show sizeable association. Finally, Díaz et al. (2012) confirms the importance of cognitive skills and personality traits as determinants of labor earnings for urban areas of Peru.[12] Specifically, various traits and combinations of traits have distinct effects on earnings: perseverance (grit) and emotional stability have a high positive influence on earnings, while agreeableness has negative effects.

The returns to higher levels of cognitive and socio-emotional skills are in some cases distinct across genders. Agreeableness and conscientiousness seem to be more rewarding for women whereas antagonism (the opposite of agreeableness), emotional stability (the opposite of neuroticism) and openness to experience were more rewarded among men in the United States (Mueller and Plug 2006). Similar results for Germany show extraversion and agreeableness negatively affect women's wages while the former has a positive effect on men's wage and the latter has no significant effect (Heineck and Anger 2010). Locus of control (self-motivation),

[10] In this study, leadership ability captures the ability to function in the very demanding environment of armed combat: willingness to assume responsibility; independence; outgoing character; persistence; emotional stability, teamwork and power of initiative (Lindqvist and Vestman 2011).
[11] Self-efficacy refers to how individuals perceive their capability to organize their work and achieve their goals (Bassi et al. 2012, p.93).
[12] Cognitive skills are measured by test scores (Peabody Picture Vocabulary, verbal fluency, working memory, and numeracy/problem-solving) and personality traits by the Big-Five model and Grit – the Duckworth Scale adapted to the Peru context (Díaz et al. 2012).

persistence and self-esteem seem to play a predicting role too on labor market outcomes but with slightly variations across gender (Heckman et al. 2006). Locus of control has also been found to be with aggression and withdrawal strong predictor of wages for white women in the United States and the United Kingdom (Osborne-Groves 2005). However, difference in the Big-Five traits and locus of control between men and women explain only modestly the gender wage gap in Australia, Germany, the Netherlands, the Russian Federation and the United States (Mueller and Plug 2006, Fortin 2008, Linz and Semykina 2008, Manning and Swaffeld 2008, Braakmann 2009, Cobb-Clark and Tan 2011). In Australia, women earn less on average not because they hold different occupations but they earn less than their men colleagues (Cobb-Clark and Tan 2011). As a consequence, the expected benefits of entering in different occupations may depend on one's abilities in ways that differ for men and women.

Cognitive and socio-emotional skills are also rewarded differently across occupations. More complex jobs, i.e. more demanding in information processing, require high-order cognitive skills (fluid intelligence). This is the case for instance of professors, scientists and senior managers (Schmidt and Hunter 2004). Conversely, conscientiousness matters for a wider spectrum of job complexity (Barrick and Mount 1991). Higher levels of socio-emotional abilities are more important for some occupations requiring low-order cognitive skills, especially in the service sector (Bowles et al. 2001b). Occupational choices are driven by personality traits such as being a caring or a direct person in adolescence (Borghans et al. 2008b). Individuals partly select occupations that correspond to their orientations. The relative price for directness over caring determines wages. Finally, traits related to Grit (persistence and motivation for long-term goals) seems to be essential for success no matter the occupation through their effect on education achievements (Duckworth et al. 2007).

The returns to skills differ across type of work as well, namely between salaried workers, incorporated self-employed and unincorporated self-employed. A mixture of traits seems to matter both for becoming and succeeding as an entrepreneur (Levine and Rubinstein 2013). Individuals with high-order cognitive skills (learning aptitudes and success as a salaried), tendency to "break-the-rules" (as measured by the degree to which the person engaged in illicit and risky activities before the age of 22) and high self-esteem in adolescence are more likely to become successful incorporated entrepreneurs in the United States (Levine and Rubinstein 2013). These abilities and traits more rewarding for incorporated entrepreneurs than for unincorporated ones and salaried workers. In the Netherlands, low-order cognitive skills and socio-emotional abilities tend to have a stronger impact on entrepreneurial incomes than on wages for salaried workers (Hartog et al. 2010). In particular, language and clerical abilities have a stronger impact on wages, whereas mathematical and technical ability as well as extraversion in early childhood are more valuable for entrepreneurs. Moreover, entrepreneurs with a balance in abilities across different fields, referred as the "Jacks-of-All-Trades", have a higher income vis-à-vis employees (Lazar 2005, Hartog et al. 2010).

Less attention has been paid by economists on the impact of technical skills on labor outcomes including wages. The role of these abilities is often associated with the impact of interventions like job trainings or vocational education rather than through the outcomes due to the acquisition of these skills (Cunningham and Villaseñor, 2014). However, a couple of studies initiated the analysis of the impact of vocational abilities of teenagers on their labor outcomes after graduation using the same longitudinal data than analogous studies on cognitive and socio-emotional skills. They show that vocational abilities (mechanical abilities, psychomotor abilities, manual dexterity and eye-hand-foot coordination) have positive effects on labor income but with considerably lower returns than to cognitive and non-cognitive abilities in the United States. However, individuals with highest level of vocational ability but low levels of standard ability (cognitive and non-cognitive) benefit from not going to college: their set of skills is associated with the highest expected hourly wage (Prada and Urzúa 2014).

However, the identification of the determinants of earnings featured in most of the previous-mentioned empirical studies faces severe measurement issues. In addition to measurement errors, simple econometric models fail to capture appropriately the effect of skills when accounting both for education and measured skills. Schooling is an endogenous variable as intelligence, personality and other factors determine it and are influenced by it (Heckman et al. 2006). Removing schooling from the wage equation overestimates the net effect of intelligence and personality on wage; meanwhile controlling for schooling in conventional specifications leads to biased, unreliable results. Beyond issues linked to measurement errors, the literature also faces the usual challenges inherent to causal analysis. First, empirical explorations of data on skills using are useful to unveil associations but do not allow to assess if variations in skills stock *cause* variations in labor outcomes and schooling decisions. Second, simple econometric models fail to capture appropriately the effect of skills. A central concern is endogeneity as certain skills are very likely to determine schooling as well as being influenced by it. Cognitive skills and education are highly correlated so distinguishing their respective effect is technically challenging, at least; an issue unaddressed in the early 2000s' literature (Cawley et al. 2001).[13]

More advanced econometric models using instrumental variables (IVs) have not yet yield convincing results given the weak predictive power of available instruments. An alternative method was adapted to this context to bypass previous sources of error: a nonparametric estimation of latent cognitive and socio-emotional abilities (using a low-dimensional vector) (Heckman et al. 2006; Urzúa 2008). It captures latent –unobservable - skills as opposed to measured skills. The former encompasses the effects measured skills (and schooling).

[13] An additional concern is to pay attention to reverse causality when analyzing the impact of neuroticism on labor market outcomes. Indeed, evidence suggests that participation to the labor market affect this trait (Gottschalk, 2005). A way to palliate to this issue is to use measures of personality traits before individuals enter the labor market (Judge and Hurst 2007, Drago, 2008).

b. The role of skills and traits on labor supply outcomes

Some of the Big-Five traits have been found to have distinct effect on labor participation, with notable gender specificities. As on earnings, conscientiousness has a large a large and positive effect on labor participation in the United States and Germany; so does extraversion (Barrick and Mount 1991, Wichert and Pohlmeier 2010). In the contrary, neuroticism and openness have a negative effect in Germany, while agreeableness only has a negative effect on labor force participation decisions of married women and no effect for other population subgroups (Wichert and Pohlmeier 2010). Socio-emotional skills have a substantial effect on the probability of employment in many, though not all, occupations in ways that differ by gender (Cobb-Clark and Tan 2011). In the United States, a man who would move from the 25th to the 75th percentile of the distribution of socio-emotional skills (locus of control and self-esteem in this case) would increase his probability of being employed at age 30 by 15 percent (Heckman et al 2006). The effects on work experience are equally important. For the case of occupational outcomes, disagreeable, intelligent individuals, achieved higher occupational status, whereas disagreeable, low intelligent men were more likely to be unemployed or working at a lower status job (Kern et al. 2013). Studies based on psychological assessments of Swedish enlistment data also show that men lower level of leadership skills have a higher probability of being unemployed than men with lower low-order cognitive abilities (Lindqvist and Vestman 2011).

Longitudinal datasets have also shown important effects of socio-emotional skills on labor outcomes. Behaviors of children in Great Britain affect significantly the probability of being in work as an adult. While hostility towards adults (at age 11) has a negative impact on the probability of being in employed at age 42, anxiety for acceptance by adults has a positive and significant impact on employment status (Carneiro et al. 2007). A potential explanation given by the study is that children who are maladjusted on this dimension are judged by their teachers to be over-zealous - which may be better rewarded in the labor market.

For the case of Latin America, like for labor earnings, in Argentina and Chile, a high-level socio-emotional skill (self-efficacy) seems to be the main determinant associated with labor force participation in both countries (Bassi et al. 2012). Patterns are almost entirely similar across the two countries. Other skills do not show strong association with labor force participation[14] except that high-order cognitive skills are also associated with higher participation in the Argentine labor force. Regarding to probability of being employed, patterns of associations with skills are similar.

[14] Skills analyzed in this study are low- and high-order cognitive skills, socio-emotional skills (such as communication, leadership and self-efficacy) (Bassi et al. 2012).

c. The role of skills on schooling decisions

While schools are privileged places to foster skills, measures of cognitive and socio-emotional skills influence schooling decisions and a range of educational outcomes (Almlund et al. 2011). Cunha et al. (2010) estimate that 12percent of the variance in educational attainment is explained by personality measures and 16percent accounted for by cognitive ability measures. Duncan et al. (2007), using longitudinal surveys of children in the United Kingdom, the United States and Canada to evaluate school readiness shows that low-order cognitive skills – in this case, mathematics, reading and attention skills - were strong predictors of later academic achievement. By contrast, measures of socio-emotional skills at school entry had limited power in explaining educational success[15].

The Big-Five traits have distinct effects on schooling. Among these traits, conscientiousness is the main determinant of overall attainment and achievement, such as college grades (Almlund et al. 2011). Openness to experience affects also educational attainment but predict attendance and the difficulty of courses selected as well. Neuroticism – as captured by self-esteem and locus of control – influences also educational attainment like graduating from a four-year college (Heckman et al. 2006). It is to be noted that the relationship of neuroticism with schooling is no always monotonic. Broader measures of personality traits influence also students' performance in test scores (Heckman et al. 2006, Almlund et al. 2011, Borghans et al. 2011, Heckman and Kautz, 2012).

Misbehavior at young age, childhood or adolescent drives lower probabilities to stay longer at school in Great Britain and the United States (Carneiro et al. 2007, DiPrete and Jennings 2011, Segal 2012). There are substantial differences between young boys and girls in their acquisition of skills from kindergarten to fifth grade. Boys and girls have roughly the same academic return to social-emotional skills but girls begin school with more advanced social and behavioral skills and their skill advantage grows over time DiPrete and Jennings 2011).

Psychology research shows that self-disciple and grit are crucial determinant of adolescent academic success in the United States. Self-discipline (on several measures) outdoes IQ as a predictor of the academic performance of adolescents: self-discipline measured in the autumn accounts for twice as much variance as IQ in explaining final grades, in final grades, high school selection, school attendance, hours spent doing homework, hours spent watching television (inversely), and the time of day students began their homework (Duckworth and Seligman 2005). Grit has recently been found to be correlated with a range of schooling success such as educational attainment, grades and retention (Duckworth et al. 2007).

Technical skills at young age can also influence the probability of going to college. By contrast to cognitive and socio-emotional skills levels, higher level of vocational ability in the United

[15] This could be explained by the fact that those measures of socio-emotional skills influence measures of cognitive skills and therefore underestimate their effect.

States is associated with lower probability to attend 4 year college (Prada 2014, Prada and Urzúa 2014)[16]. This may not be surprising given that individuals with high vocational abilities are likely to be good at performing manual tasks (Autor et al. 2003); and thus they are likely to enter in vocational programs rather than to college. In fact, it may not be only an orientation toward adequate with an individual's skills set but also an expectation for higher labor earnings. Given that individuals in the top 10percent of vocational skills have higher hourly wages than those in the 10 percent of cognitive and socio-emotional abilities, not going to college is indeed associated with higher expected earnings.

4. Data

The analysis in this paper is based on the Skills Toward Employment and Productivity (STEP) Household Survey, a multi-country study led by the World Bank assessing skills that matter and their links to the labor market. The STEP household survey of Colombia is representative of the country's thirteen main cities and their metropolitan areas (see Appendix for additional information). This covers the large majority of Colombia's urban population and is the area widely used by labor market household surveys in Colombia. The sample size is 2,617. The distribution in age, gender and education is similar than national household surveys for the same urban areas.

The STEP household survey elicits a wide range of information on personal background, education, employment and compensation, household wealth, and household size and composition, similar to a standard household survey. Additionally, a randomly-selected individual in each household is further survey on information related to reading proficiency, personality, personal health, and technical skills used on the-job. The measures of skills include:

Cognitive skills measures: Survey respondents take a reading proficiency test developed by the Educational Testing Service (ETS), drawing on ETS' work for the Program for the International Assessment of Adult Competencies (PIAAC). The test consists of several sub-tests: Reading Components that assesses "foundational reading skills", namely word meaning, sentence processing and reading comprehension; core literacy, and, for those who "pass" the first two tests, four advanced reading tests. Rather than generating a reading proficiency test value based on each participant's measured scores, a set of 10 "plausible values" is generated for each respondent based on post-collection estimation relying on all available background information (gender, age, metropolitan area of living, etc.). Each estimation featured in this paper is repeated 10 times for each plausible value. The average coefficients and standards errors of the 10 estimations are reported (Von Davier et al. 2009, OECD 2013).

[16] Measures of vocational abilities are constructed from three sections of the Armed Service Vocational Aptitude Battery (ASVAB): auto and shop information, mechanical comprehension and electronics information. The dataset used is the National Longitudinal Survey of Youth (NLSY79).

Socio-emotional skills measures: The battery is composed of: the short Big Five Inventory, consisting of fifteen items; a seven-item risk and time preference scale; and three items to capture grit, two to capture hostile attribution bias, and four to capture decision making skills.

5. Empirical strategy

a. Conventional approach

We would like to estimate the following relationship between labor market outcomes and skills:

$$Y_i = \alpha + \beta_1 A_i + \beta_2 X_i + \varepsilon_i \qquad (1)$$

where Y_i is a labor market outcome, A_i represents all ability (skills) that affects the labor market outcome, and X_i is a set of factors (other than ability) that affect Y_i. Since ability is unobserved, the measure commonly used for A_i is years of schooling or school levels. Assuming that highest year of schooling completed S_i captures all A_i, we can substitute out A_i for S_i which, if Y_i is wages, gives us the typical Mincerian wage equation:

$$Y_i = \alpha + \beta_1 S_i + \beta_2 X_i + \varepsilon_i \qquad (2)$$

If the decision-maker for our outcomes of interest – wages, employment, occupation – is the employer, S_i may be a perfect measure for the ability information that the employer has at the time of hiring. Employers receive curricula vitae that list educational attainment and little else. Thus, that employer is making hiring and/or wage decisions based on the same information that the researcher has. The estimated β_1 will be an unbiased estimate of the impact of ability on labor market outcomes.

However, employers are likely to have more information than schooling attainment, such as that obtained through an interview, where the employer assesses communications skills, professional behavior, and other non-cognitive skills. In other words, S_i is unlikely to capture all ability measurement[17]. Thus, S_i is correlated with ε_i, which results in biased estimates of the returns to schooling (β_1). Suppose that T_i measures all skills that are captured in S_i and ε_i in equation (2). We can rewrite equation (2) as

$$Y_i = \alpha + \beta_1 T_i + \beta_2 X_i + \upsilon_i \qquad (3)$$

Under the assumption that our set of T_i perfectly measures all A_i,[18] we can estimate equation (3) using OLS without any ability bias and β_1 will give us the return to each skill captured by vector T_i. Returning to our employer's information, if she only has information about school test

[17] Educational attainment is a poor measurement of actual scholastic ability, as shown by the high variance in grade-specific achievement tests that are administered in various countries.

[18] We are careful to define T_i as "measured skills" and not as skills. Heckman et al. (2006) posit a latent variable that affects measured skills (and schooling); we adopt that assumption, as well, as recommended by the psychology literature (Borghans et al. 2008).

scores - a proxy for cognitive skills – and behavioral information from an interview, the T_i in equation 3 may be adequate to estimate β_1 without error.

However, a growing literature shows that T_i does not fully capture A_i so $Corr(T_i, \upsilon_i) \neq 0$. In that case, we again have a biased estimate of β_1.[19] The literature has addressed this problem through various means. One approach is to use an instrumental variable, Z_i that is correlated with T_i but not correlated with υ_i. The weaknesses of this approach are multiple, including difficulty in selecting a good IV for T_i, the need to select an IV for each T_i (many of which are produced through similar processes), and the inability of IV to solve measurement error. For a discussion of these challenges, see Heckman et al. (2006) and the sources cited therein.

b. Structural Estimation

An alternative to OLS and IV estimation as exposed before is to consider the following reduced-form equation:

$$Y = X_Y\beta^Y + \alpha_A^Y\theta_A + \alpha_B^Y\theta_B + e^Y \qquad (5)$$

Where Y is the outcome of interest (e.g., wage), X_Y are observable controls (e.g., gender, age), θ_A, θ_B are the latent factors or dimensions of unobserved heterogeneity, β^Y, α_A and α_B^Y are coefficients to estimate, and e^Y is a vector of independently distributed error terms orthogonal to X_Y, θ_A and θ_B. The need of a structural estimation relies on the assumption that θ_A and θ_B are unobservable. That is, the measures or scores we have in the data are only proxies of the true latent variables that we want to use for the estimation (Bartholomew et al., 2011). They are treated as realizations of a score-production function, presented in (6), whose inputs are observable and unobservable characteristics.

$$T = X_T\beta^T + \alpha_A^T\theta_A + \alpha_B^T\theta_B + e^T \qquad (6)$$

Where T is a $L \times 1$ the vector of scores (e.g., measures of IQ, coding speed, self-esteem or locus of control), X_T is a matrix of observable controls and e^T is a vector of independently distributed error terms orthogonal to X_Y, θ_A, θ_B and e^Y. In this sense, the model comprises a measurement system (i.e., outcomes, test scores, observable controls and error terms) that is linked by latent factors or unobserved heterogeneity (i.e., θ_A and θ_B). In this case, the identification assumption assumes that e_Y and e_T are mutually independent conditional on (θ, X). This is the type of models

[19] Test scores are sensitive to the amount of schooling completed at the time of the test and family background. Further, the measures of ability are known to be very noisy. Thus, using test scores as an independent variable in regression model analysis will lead to measurement error bias.

considered by Keane and Wolpin (1997), Cameron and Heckman (2001), Heckman et al. (2006), Urzúa (2008), Sarzosa and Urzúa (2013, 2014).

Carneiro et al. (2003) based on the findings of Kotlaski (1967) showed that we can use the system of test scores (6) and their production function to non-parametrically identify the distributions of the latent endowments $f_{\theta_A}(\cdot)$ and $f_{\theta_B}(\cdot)$, their coefficients (α_A and α_B, up to one normalization, that is, we need to set one coefficient equal to one and the rest will be interpreted relative to the one chosen as nummeraire) and the diagonal matrix of their variance Σ_θ with the help of two restrictions.[20] First, θ_A and θ_B need to be orthogonal to each other, and that we need to have at least three test scores per factor.[21] That is, if we have two factors, $L = 6$. In practice, the test scores measurement system allows us to identify the distributions that are followed by the unobserved heterogeneity, in order to be able to integrate it away in a Maximum Likelihood procedure.[22] The likelihood function is then:

$$\mathcal{L} = \prod_{i=1}^{N} \int \int f_{e^Y}(X_Y, Y, \varrho_1, \varrho_2) \times f_{e^{T_1}}(X_{T_1}, T_1, \varrho_1, \varrho_2) \cdots \times f_{e^{T_6}}(X_{T_6}, T_6, \varrho_1, \varrho_2) dF_{\theta_1}(\varrho_1) dF_{\theta_2}(\varrho_2) \tag{7}$$

Maximizing (7) we can retrieve all the parameters of interest: β_Y, β_{T_τ}, α_A^Y, α_B^Y, $\alpha_A^{T_\tau}$, $\alpha_B^{T_\tau}$ for $\tau = \{1,2,3,4,5,6\}$, and the parameters (i.e., the means, standard deviations and mixing probabilities) that describe the distributions $f_{\theta_A}(\cdot)$ and $f_{\theta_B}(\cdot)$.

6. Descriptive Statistics

This section presents some descriptive statistics on the distribution of skills across gender, age groups, and educational level of the working-age population.

With regards to reading proficiency, Figure 1 shows that while there seems to be almost no difference by gender, the distribution of scores is highly correlated (as expected) with educational level, though not perfectly. In particular, the difference between those with secondary and tertiary education is not that pronounced as that between those with primary and secondary education. Also interestingly, among youngster (15-24) the distribution of scores is higher than among adults, a signal that suggests improvement over generations in the ability to

[20] It should be noted that the estimated distributions $f_{\theta_A}(\cdot)$ and $f_{\theta_B}(\cdot)$ are not assumed to follow any particular distribution. The procedure uses a mixture of normals, which are known to be able to recreate a wide range of distributions (Frühwirth-Schnatter, 2006).

[21] Carneiro et al. (2003) actually propose a slightly milder condition: $L > 2k + 1$, where k is the number of latent factors in the system.

[22] Integrals are calculated using the Gauss-Hermite quadrature (Judd, 1998).

read (also possibly correlated with educational levels) or that this ability tends to depreciate with aging.

In terms of socio-emotional skills (Figures 2 to 4), differences across gender, age, and educational level are less noticeable. Across gender, and among all possible dimensions covered in the survey, men and women have slight differences in distribution of conscientiousness, emotional stability, grit, decision making and hostile attribution bias scores. The most noticeable is the fact that males tend to score higher in the emotional stability scale than women. By age, the only differences are registered for agreeableness, emotional stability, grit and hostile attribution bias, with youngsters scoring lower than adults across these dimensions. Finally, there are significant differences in socio-emotional skill scores by educational level: in all cases except for hostile attribution bias (which is the opposite), less educated workers score lower in the scale.

It is important to remark that these distributions are not conditional on other observable and unobservable characteristics of individuals.

We also explore the extent of correlation among cognitive and socio-emotional skills, as well as within different dimensions of it. As shown in Table 1, the correlation between reading proficiency and socio-emotional dimensions differ substantially, with openness to experience, decision making, and hostile attribution bias among the ones with higher correlation, but never higher than 0.25. Some socio-emotional dimensions are also relatively higher correlated among themselves: for instance, extraversion with openness to experience (0.17); emotional stability with hostile attribution bias (-0.17); conscientiousness with grit (0.21), decision-making (0.17), agreeableness (0.16), and openness to experience (0.16); openness to experience with decision making (0.29), agreeableness (0.20). and grit (0.20); agreeableness with grit (0.21), and decision making (0.17); and decision making and grit (0.21).

7. Results

a. OLS Estimates

Our first set of results explores OLS estimates on the relationship between cognitive and socio-emotional skills and labor market outcomes. The first set out outcomes concern log hourly labor earnings (wage for salaried workers, and net profit for self-employed).These results are presented in Table 1. Sample include individuals between 15 and 64 years of age (both males and females).

The main conclusion is that, controlling for other observable characteristics such as gender, age, mother's education, and regional indicators, reading proficiency is positive and statistically significantly related to wages. However, for the case of socio-emotional skills, only openness to experience seems to be significantly related to wages. These results remain when all skill dimensions are included in the same regression. It is important to notice that estimates in Table 2 do not control for educational level, so the interpretation of the coefficients is that they capture the full association between different skills dimensions and wages, irrespectively of where these skills were formed (at school, at work, at home, etc.).

Table 3 shows results for other labor market outcomes and occupational choices, including the likelihood of being a formal worker, of being a white-collar worker, of being employed, of being active or in school, or having attended tertiary education. Reading proficiency is again positively related with the probability of being formal or white-collar workers, but socio-emotional skills seem to play no role in these outcomes (the exception seems to be being hostile and being a formal worker). However, some of these socio-emotional characteristics seem to be relevant for labor or education choice paths. For instance, conscientiousness and decision making are positively related to being employed, looking for a job, or in school. And a higher scale in openness to experience, emotional stability, decision making, and hostile attribution bias seem to matter for attending tertiary education. For comparison purposes, Table 3 also includes regressions that control for educational level of the individual, with the only difference being in the role of reading proficiency, which becomes non-significant, suggesting that this skill is mainly being formed at school (as mentioned in the section before, both are highly correlated).

Table 4 presents results for different gender, age, and education groups. It only shows for two outcomes, namely hourly earnings, and being active or in school, but other results for different labor outcomes are available upon request. The main findings show that reading proficiency (without controlling for education) remains positive and statistically significant in relation to wages across gender and age, but only among the more educated individuals (with at least 9 years of education). In contrast, it is only related with labor force or school participation among females, youngsters (less than 35 years old), and less educated workers. Regarding socio-emotional skills, the role of openness to experience on wages seem only relevant among males, older, and more educated workers. The role of socio-emotional skills on explaining labor and schooling decision also affects differently across subgroups. Strikingly, among males, socio-emotional skills do not play any role in occupational decisions.

b. IV Estimates

As described in Section 5.a, an alternative to OLS is to do IVs to attempt estimates free of measurement error or omitted variable bias. For that, it is needed to find a set of instruments Z_i that correlates with the skills set T_i in equation (3), but not with the error term υ_i.

Among the possible instruments that could be appropriate and available in the Colombia STEP survey, we selected them on the basis of compliance with the over-identifying restriction test (Sargan-Hansen test). We came up with the age at which a person started school, and the economic situation of the household at age 12, as suitable instruments for reading proficiency score. And for each socio-emotional skill, we use the indicator of whether the individual lived with both parents at age 12, and again the economic situation of household at age. 12. These instruments are valid per Sargan-Hansen test.

Table 5 presents the main results of IV estimation (second stage) for each labor market outcome, considering just on skill dimension at a time in each regression (for presentational purposes, the coefficients from different regressions for same outcome are presented in one column). Most of the same results found using OLS remain, in particular the fact that only reading proficiency seems to matter for explaining earning, rather than socio-emotional skills. No set of skill is significant for the probability of being a formal worker, of being employed, or being active or in school. In contrast, reading proficiency and several socio emotional skills (conscientiousness, emotional stability, grit, hostile attribution bias) seems to be significantly related with the probability of being a white-collar worker, or having attended tertiary school.

c. Structural Estimation

As stated in Section 5.b, we can also estimate the effect of latent skills on labor market outcomes using structural estimation methods developed as in Keane and Wolpin (1997), Heckman et al. (2006) and Sarzosa and Urzúa (2014). In this particular context, it is important to notice that since the STEP Colombia survey is a cross-sectional data set, we measure the latent traits at the same time we measure the outcome variables. Therefore, this method would not solve the reverse causality concern, as some of the scores retrieved from latent skills might still be affected by employment or education choices. However, as opposed to the OLS and IV estimates presented before, this method can mitigate measurement error concern since it relies on latent skills that fully capture their relation with outcomes abstracting from single (and potential poor) measures of skills and traits.

To apply this method, we construct an adjunct measurement system that comprises scores in two dimensions: reading proficiency and socio-emotional skills. Carneiro et al. (2003) show that identification in this set-up requires at least three test scores per dimension explored. Therefore, we need to construct three scores that provided information about socio-emotional skills and three scores that provided information about reading proficiency.

In order to obtain the former, and following the correlations founds and described in Table 1, we aggregate the scale of extraversion with the measure openness to experience in one score, the measure of emotional stability with the measure of hostile attribution bias in another score, and measures of conscientiousness, grit and decision-making in the third score.[23] To obtain the measurement system needed to identify reading proficiency we used the following scores: i) a constructed measure of reading components that aggregated –through a weighted average–

[23] Aggregation of these measures was needed in order to secure the necessary smoothness in the measurement system, as all of these measures come from categorical answers. The pairing of the measures was done based on the correlations among them.

scores of Average Print Vocabulary (AVP), sentence processing and passage comprehension; ii) a plausible value from the Core Literacy test; and iii) a measure of use and length of the reading done on and off the workplace.

Using these scores and exogenous controls like age, gender, mother's education and city in which they live, we can estimate the system of equations described in equations (5) and (6).[24] The purpose of these estimations is to retrieve the components of the unobserved heterogeneity clean from exogenous characteristics that can affect the scores we observe. These estimations are presented in Tables 6 and 7. For instance, people with more educated mothers are more likely to have a broader vocabulary and hence score higher in the AVP even if the latent reading proficiency is unchanged.

But more important that these coefficients are the estimated distributions of the unobserved heterogeneity obtained from these estimations. These distributions will be used to structurally model the unobserved heterogeneity in the outcome equations. Figure 5 presents the variance decompositions of the scores. We see that the latent factors explain big proportions of the variance of many of the scores. That clean variation is the one we identify as the latent skill or unobserved heterogeneity.

Having estimated the distributions that describe reading proficiency and socio-emotional skills, we estimate their relation with labor market outcomes. The results presented in Table 8 indicate that the unobserved heterogeneity matters in almost every outcome we analyzed, yet in very different ways. Socio-emotional skills matter in choices like participating in the labor market and attending to college. That is, socio-emotional skills prevent inactivity both among working and studying aged population, although the probability of attending to college is also highly correlated with reading proficiency. Once in the labor market, reading proficiency is the one related with higher probabilities of working in the formal sector, being a white-collar worker and earning more. In fact our results indicate that an increase in one standard deviation in reading proficiency is associated with an increase of 12.5 percent in earned hourly wages.

Given the unobserved nature of our traits of interest, we must rely on simulations in order to interpret our results and better describe the size of the relations we inquire about. We present the expected outcome as a function of the unobserved heterogeneity. Therefore, given that we estimated two dimensions of such heterogeneity, we present three-dimensional graphs that represent

$$E[Y|\theta_A, \theta_B] = E[X\beta] + \alpha_A\theta_A + \alpha_B\theta_B \tag{8}$$

[24] All the estimations presented in this chapter were implemented using the heterofactor command in Stata developed in Sarzosa and Urzúa (2014).

In that sense, we randomly draw θ_A and θ_B from the distributions estimated in the first step estimations (described in Tables 6 and 7) and construct $E[Y|\theta_A, \theta_B]$ based on the draws, the exogenous controls and the estimates for β, α_A and α_B. This way, we clearly see how the unobserved skills relate with the outcome variable.[25] Interpretation is also aided by the fact that the scale of the unobserved heterogeneity is presented in terms of deciles of their respective distribution.

Figure 6 shows that the probability of being active either in the labor force or as a student raises as the socio-emotional and the reading proficiency increase. While low skilled population has a 78percent probability of being active, high skilled population has a 95 percent probability of being so. If focused on only one dimension, we see that ceteris paribus, a person gains 9 percentage points in the probability of being active if taken from the first decile to the tenth decile in the personality traits distribution. In the same way, there is an increase of 5.8 percentage points in the probability of being active associated with taking a person from the first decile to the tenth decile of the reading proficiency distribution.

Figure 7 shows that the relation between going to college and skills is even stronger. While those with the least skills have almost no chance of going or having gone to college (only 1.5 percent), those with the highest levels of skills have an 83 percent probability. Although both set of skills are correlated with this outcome, the size of such relation is dramatically different. Changing a person's personality traits from the first decile to the tenth decile of the distribution is associated with an increase in the probability of going or having gone to college by 17.7 percentage points. The increase rises to 71.2 percentage points when we compare those in the first decile with those of the tenth decile of the reading proficiency distribution, leaving everything else constant.

The size of these relations contrasts with the ones that arise when we analyze the probability of being employed. Figure 8 shows that the probability of being employed remains unchanged at around 75 percent in the entire skills space. However, once the decision of working is out of the way, the quality of the job does correlate with skills, in particular, with reading proficiency. Figures 9 to 11 attest to this. For instance, Figure 9 shows that all else constant, the likelihood of having a formal job increases in 28 percentage points (i.e., more than doubles) when we compared a person in the first decile with that of a tenth decile of the reading proficiency distribution. In the same way, high skilled workers in terms of reading proficiency are 26.7 percentage points more likely to have a white-collar job than low skilled workers who have a 28 percent probability of doing so (Figure 10). Also, workers with highest score in reading proficiency can earn up to CO$ 3,000 (equivalent to USD 1.5) more per hour than those in the lowest scale, which is roughly 50 percent more (Figure 11).

[25] For the case probit case the expected outcome equation follows the same logic. Therefore it becomes: $E[Y|\theta_A, \theta_B] = \Pr(E[X\beta] + \alpha_A\theta_A + \alpha_B\theta_B + \zeta > 0)$ where $\zeta \sim \mathcal{N}(0,1)$.

8. Conclusion

Using a unique dataset that measures cognitive (reading proficiency) and socio-emotional skills (personality traits and behavior) for Colombia, this paper has documented the role that these skills play in the labor market by looking at different outcomes for different subsets of the population, and different methodologies (OLS, IVs, Structural estimation).

Across all methods, one result came up quite consistently: reading proficiency is an important predictor of earnings and quality job. For instance, using our preferred methodology, one standard deviation in the scale of reading proficiency can increase hourly wages by 12.5 percent. Reading proficiency is also an important predictor of being a formal or white-collar worker. This result is consistent with previous literature findings, for instance Murnane et al. (1995), Murnane et al. (2000), Altonji and Pierret, (2001), Cawley et al. (2001), and Hanushek and Woessmann, (2008) for the US.

In contrast, the role of socio-emotional skill is more mixed. Across all methodologies explored, they do not seem to play any significant role in explaining wage levels, or job quality. This finding is at odds with previous literature for the US, such as Bowles et al. (2001a, 2001b) and Drago (2011). But they do seem to play an important role as predictor of labor force participation and schooling decisions, along with cognitive skills, as found in Carneiro et al. (2007) and Almlund et al. (2011).

Some results emerge as differential for different subgroups (due to data limitations, it was only possible to perform this breakdown using OLS methods). For instance, socio-emotional skills are more important predictor of labor force participation among women, the younger (under 35), and less educated workers (less than complete secondary education). In contrast, cognitive skills seem more relevant for explaining wage levels among males, older, and more educated workers.

These results have important policy implications for school and vocational training programs in terms of curricula, where the combination of development modules of cognitive and socio-emotional skills would play quite distinctive roles, depending on the immediate policy objective, namely improving job quality or fostering higher labor market participation or tertiary education. In any case, further research is needed on the optimal combination of packages for different demographic and socio-economic population groups.

9. References

Aedo, C. and I. Walker (2012), *Skills for the 21st Century in Latin America and the Caribbean*, World Bank, Washington, DC.

Almlund, M., A. L. Duckworth, J. J. Heckman and T. Kautz (2011), “Personality Psychology and Economics”, in E. A. Hanushek (ed.), *Handbook of the Economics of Education*, Volume 4, pp. 1-181, North Holland, Amsterdam.

Altonji, J. G. and C. R. Pierret (2001), “Employer Learning and Statistical Discrimination”, *Quarterly Journal of Economics*, Vol. 116, No. 1, pp. 313–50.

Autor, D. H., H. Levy and F. R. Murnane (2003), “The Skill Content of Recent Technological Change: An Empirical Exploration”, *Quarterly Journal of Economics*, Vol. 118, No. 4, pp. 1279-1334.

Banerji, A., W. Cunningham, A. Fiszbein, E. King, H. Patrinos, D. Robalino and J.-P. Tan (2010), *Stepping Up Skills for More Jobs and Higher Productivity*, The World Bank, Washington, DC.

Barrick, M. R. and M. K. Mount (1991), “The Big Five Personality Dimensions and Job Performance: A Meta-Analysis”, Personnel Psychology, Vol. 44, No. 1, pp. 1–26.

Bassi, M., M. Busso, S. Urzúa and J. Vargas (2012), *Disconnected: Skills, Education and Employment in Latin America*, Inter-American Development Bank, Washington, DC.

Bassi, M. and S. Urzúa (2010), "Educación en Chile: El Desafío está en la Calidad" (Education in Chile: The Challenge is the Quality), *IDB Technical Note*, No. 271, Inter-American Development Bank, Washington, DC.

Borghans, L., A. L. Duckworth, J. J. Heckman and B. ter Weel (2008a). “The Economics and Psychology of Personality Traits,” *Journal of Human Resources*, Vol. 34, No. 4, pp. 972-1059, University of Wisconsin Press.

Borghans, L., B. H. H. Golsteyn, J. J. Heckman and J. E. Humphries (2011), “Identification Problems in Personality Psychology”, *Personality and Individual Differences*, Vol. 51, No. 3, pp. 315–320.

Borghans, L., B. ter Weel and B. A. Weinberg J. (2008b), “Interpersonal Styles and Labor Market Outcomes”, *Human Resources*, Vol. 43, No. 4, pp. 815-858.

Bowles, S., H. Gintis and M. Osborne (2001a), “Incentive-Enhancing Preferences: Personality, Behavior, and Earnings”, *American Economic Review*, Vol. 91, No. 2, pp.155–158.

Bowles, S., H. Gintis and M. Osborne (2001b), "The Determinants of Earnings: A Behavioral Approach", *Journal of Economic Literature*, Vol. 39, No. 4, pp. 1137-1176.

Braakmann, N. (2009), “The Role of Psychological Traits for the Gender Wage Gap in Full-time Employment and Wages: Evidence from Germany”, *SOEP papers on Multidisciplinary Panel Data Research Paper*, No. 162.

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

Carneiro, P., K. Hansen and J. J. Heckman (2003), ‘‘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*, Vol. 44, No. 2, pp. 361–422.

Castro, J. F., and G. Yamada (2012) “‘Convexification’ and ‘Deconvexification’ of the Peruvian Wage Profile: A Tale of Declining Education Quality”, *Working Paper DD/12/02*, Universidad del Pacífico, Lima.

Castro, J. F., G. Yamada and O. Arias (2011). “Higher Education Decisions in Peru: On the Role of Financial Constraints, Skills, and Family Background”, *Working Paper DD/11/14*, Universidad del Pacifico, Lima.

Cattell, R. B. (1987), *Intelligence: Its Structure, Growth, and Action*, Elsevier Science Pub. Co., New York, NY.

Cawley, J., J. J. Heckman and E. Vytlacil (2001), “Three Observations on Wages and Measured Cognitive Ability”, *Labour Economics*, Vol. 8, pp. 419–442, Elsevier.

Cobb-Clark, D. A. and M. Tan (2011), “Noncognitive Skills, Occupational Attainment, and Relative Wages”, *Labour Economics*, Vol. 18, No. 1, pp. 1-13.

Cunha, F., J. J. Heckman, and S. M. Schennach (2010), “Estimating the Technology of Cognitive and Noncognitive Skill Formation”, *Econometrica*, Vol. 78, No. 4, pp. 883–931.

Cunningham, W. and P. Villaseñor (2014), “Employer Voices, Employer Demands, and Implications for Public Skills Development Policy”, *Policy Research Working Paper*, No.6853, The World Bank, Washington, DC.

Díaz, J. J., O. Arias, and D. V. Tudela (2012). “Does Perseverance Pay as Much as Being Smart?: The Returns to Cognitive and Non-Cognitive Skills in Urban Peru”, *mimeo*, The World Bank, Washington, DC.

DiPrete, T.A. and J.L. Jennings (2011), "Social and Behavioral Skills and the Gender Gap in Early Educational Achievement", Vol. 41, No. 1, pp. 1–15.

Drago, F. (2011), "Self-Esteem and Earnings", *Journal of Economic Psychology*, Vol. 32, No. 3, pp. 480-488.

Duckworth, A. L. and M. E. P. Seligman (2005), "Self-Discipline Outdoes IQ in Predicting Academic Performance of Adolescents", *Psychological Science*, Vol. 16, No. 12, pp. 939-944.

Duckworth, A., C. Peterson, M. Matthews and D. Kelly (2007). "Grit: Perseverance and Passion for Long Term Goals", *Journal of Personality and Social Psychology*, Vol. 92, No. 6, pp. 1087-1101.

Duncan, G. J., C. J. Dowsett, A. Claessens, K. Magnuson, A. C. Huston, P. Klebanov, L. S. Pagani, L. Feinstein, M. Engel, J. Brooks-Gunn, H. Sexton, K. Duckworth and C. Japel (2007), "School Readiness and Later Achievement", *Developmental Psychology*, Vol. 43, No. 6, pp. 1428-46.

Gasparini, L., G. Cruces, S. Galiani, and P. Acosta (2011), "Educational Upgrading and Returns to Skills in Latin America: Evidence from a Supply-Demand Framework for the Decades of 1990 and 2000", Background paper for C. Aedo and I. Walker (2012), *Skills for the 21st Century in Latin America and the Caribbean*, the World Bank, Washington, DC.

Gottfredson, L. S. (1997), "Why g Matters - The Complexity of Everyday Life", Intelligence, Vol. 24, No.1, pp. 79- 132.

Gottschalk, P. (2005), "Can Work Alter Welfare Recipients' Beliefs?", *Journal of Policy Analysis and Management*, Vol. 24, No. 3, pp. 485–498.

Guerra, N., K. Modecki, and W. Cunningham (2014), *Social-Emotional Skills Development Across the Life Span: PRACTICE*, The World Bank, Washington, DC.

Finnie, R. and R. Meng (2001), "Minorities, Cognitive Skills, and Incomes of Canadians", Canadian Public Policy, Vol. 28, No. 2, pp. 257-73.

Fortin, N. M. (2008), "The Gender Wage Gap among Young Adults in the United States: The Importance of Money vs. People", *Journal of Human Resources*, Vol. 43, No. 4, pp. 886-920.

Hanushek, E. A. and L. Woessmann (2008), "The Role of Cognitive Skills in Economic Development", *Journal of Economic Literature*, Vol. 46, No. 3, pp. 607–668.

Hartog, J., M. van Praag and J. van der Sluis (2010), “If You Are So Smart, Why Aren’t You an Entrepreneur? Returns to Cognitive and Social Ability: Entrepreneurs versus Employees”, *Journal of Economics & Management Strategy*, Vol. 19, No. 4, pp. 947-989, Wiley Blackwell.

Heckman, J. J. and T. Kautz (2012), “Hard Evidence on Soft Skills”, *Labour Economics*, Vol. 19, No. 4, pp. 451–464.

Heckman, J. J. and Y. Rubinstein (2001), “The Importance of Noncognitive Skills: Lessons from the GED Testing Program”, *American Economic Review*, Vol. 91, No. 2, pp. 145–49.

Heckman, J. J., J. Stixrud and S. Urzúa (2006), “The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior”, *Journal of Labor Economics*, Vol. 24, No. 3, pp. 411–482.

Heineck, G. and S. Anger (2010), “The Returns to Cognitive Abilities and Personality Traits in Germany”, *Labour Economics*, Vol. 17, No. 3, pp. 535–46.

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

Horn, J. L. (1970), “Organization of Data on Life-Span Development of Human Abilities”, in L. R. Goulet and P. B. Baltes (eds.), *Life-Span Developmental Psychology: Research and Theory*, pp. 423–66, Academic Press, New York, NY.

Judge, T. A. and C. Hurst (2007), “Capitalizing on One’s Advantages: Role of Core Self-Evaluations”, *Journal of Applied Psychology*, Vol. 92, No. 5, pp. 1212–1227.

Kern, M. L., A. Duckworth, S. Urzúa, R. Loeber, M. Stouthamer-Loeber and D. Lynam (2013), “Do as You’re Told! Facets of Agreeableness and Early Adult Outcomes for Inner-City Boys”, *Journal of Research in Personality*, Vol. 47, No. 6, pp. 795-799.

Kuhn, P. and C. Weinberger (2005), “Leadership Skills and Wages”, *Journal of Labor Economics*, Vol. 23, No. 3, pp. 395-436.

Kniesner, T. J. and B. ter Weel (eds.) (2008), “Special Issue on Noncognitive Skills and Their Development”, *Journal of Human Resources*, Vol. 43, No. 4, pp. 729-1059, The University of Wisconsin Press, Madison, WI.

Lazear, E. P. (2003), “Teacher Incentives,” *Swedish Economic Policy Review*, Vol. 10, No. 2, pp. 179-214.

Levine, R. and Y. Rubinstein (2013), “Smart and Illicit: Who Becomes an Entrepreneur and Does it Pay?”, *NBER Working Paper Series*, No. 19276, National Bureau of Economic Research (NBER), Cambridge, MA.

Levy, S. and N. Schady (2013) "Latin America's Social Policy Challenge: Education, Social Insurance, Redistribution", *Journal of Economic Perspectives*, Vol. 27, No. 2, pp. 193-218.

Lindqvist, E. and R. Vestman (2011), "The Labor Market Returns to Cognitive and Noncognitive Ability: Evidence from the Swedish Enlistment", *American Economic Journal: Applied Economics*, Vol. 3, No. 1, pp. 101-128.

Linz, S. J. and A. Semykina (2008), "Attitudes and Performance: An Analysis of Russian workers", *The Journal of Socio-Economics*, Vol. 37, No. 2, pp. 694-717.

Lustig, N., L. F. Lopez-Calva and E. Ortiz-Juarez, "Deconstructing the Decline in Inequality in Latin America", *World Bank Policy Research Working Paper Series*, No. 6552, The World Bank, Washington, DC.

McIntosh, S. and A. Vignoles (2001), "Measuring and Assessing the Impact of Basic Skills on Labour Market Outcomes", *Oxford Economic Papers*, Vol. 53, No. 3, pp. 435-81.

Manning, A. and J. Swaffeld (2008), "The Gender Gap in Early-Career Wage Growth", Economic Journal, Vol. 118, No. 530, pp. 983-1024.

Mourshed, M., D. Farrell and D. Barton (2012), *Education to employment: Designing a system that works*, McKinsey Center for Government, McKinsey & Company, Washington, DC.

Mueller, G. and E. J.S. Plug (2006), "Estimating the Effect of Personality on Male and Female Earnings", *Industrial and Labor Relations Review*, Vol. 60, No. 1, pp. 3-22. Cornell University, Ithaca, NY.

Mulligan, C. (1999), "Galton versus the Human Capital Approach to Inheritance", *Journal of Political Economy*, Vol. 107, No. 6, pp. 184-224.

Murnane, R., J. Willett, Y. Duhaldeborde, and J. H. Tyler (2000), "How Important are the Cognitive Skills of Teenagers in Predicting Subsequent Earnings", *Journal of Policy Analysis and Management*, Vol. 19, No. 4, pp. 547-68.

Murnane, R. J., J. B. Willett and F. Levy (1995), "The Growing Importance of Cognitive Skills in Wage Determination", *The Review of Economics and Statistics*, Vol. 77, No. 2, pp. 251-266, MIT Press.

Neisser, U., G. Boodoo, T. J. Bouchard, A. W. Boykin, N. Brody, S. J. Ceci, D. F. Halpern, J. C. Loehlin, R. Perloff, R. J. Sternberg and S. Urbina (1996), "Intelligence: Knowns and Unknowns", *American Psychologist*, Vol. 51, No. 2, pp. 77-101.

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

OECD (2013), *OECD Skills Outlook 2013: First Results from the Survey of Adult Skills*, OECD Publishing, Paris.

Osborne-Groves, M. (2005): "How Important is Your Personality? Labor Market Returns to Personality for Women in the US and UK," *Journal of Economic Psychology*, Vol. 26, No. 6, pp. 827–841.

Prada, M. (2014), "Beyond Smart and Sociable: Rethinking the Role of Abilities on Occupational Choices and Wages", *mimeo*, University of Maryland, MD.

Prada, M. and S. Urzúa (2014), "One Size Does Not Fit All: The Role of Vocational Ability on College Attendance and Labor Market Outcomes", *mimeo,* University of Maryland, MD.

Sarzosa, M. and S. Urzúa (2013), "Bullying and Cyberbullying in Teenagers - The Role of Cognitive and Non-Cognitive Skills", *paper presented at the LACEA Labor Network conference on Skills, Education and Labor Market Outcomes,* May 2-3, 2013, University of Maryland, MD.

Schmidt, F. L. and J. Hunter (2004), "General Mental Ability in the World of Work: Occupational Attainment and Job Performance", *Journal of Personality and Social Psychology,* Vol. 86, No. 1, pp. 162–173.

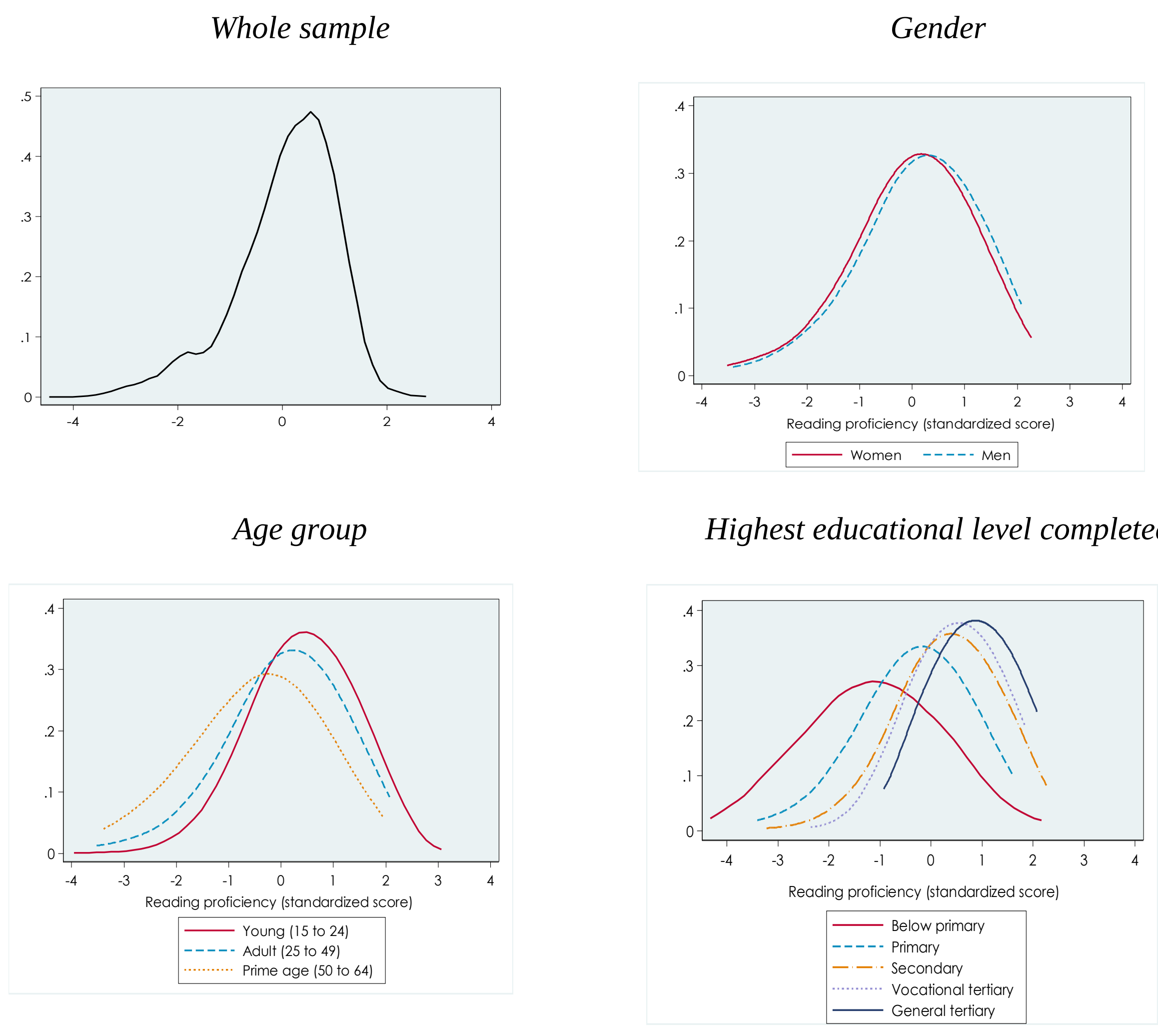
Segal, C. (2013), "Misbehavior, Education, and Labor Market Outcomes", *Journal of the European Economic Association*, Vol. 11, No. 4, pp. 743–779.

Urzúa, S. (2008), ‘‘Racial Labor Market Gaps: The Role of Abilities and Schooling Choices’’, *Journal of Human Resources*, Vol. 43, No. 4, pp. 919-971.

Wichert, L. and W. Pohlmeier (2010), "Female Labor Force Participation and the Big Five", ZEW Discussion Papers 10-003, ZEW (Zentrum für Europäische Wirtschaftsforschung / Center for European Economic Research), Mannheim, Germany.

Figure 1. Distribution of reading proficiency across groups of interest

Kernel density

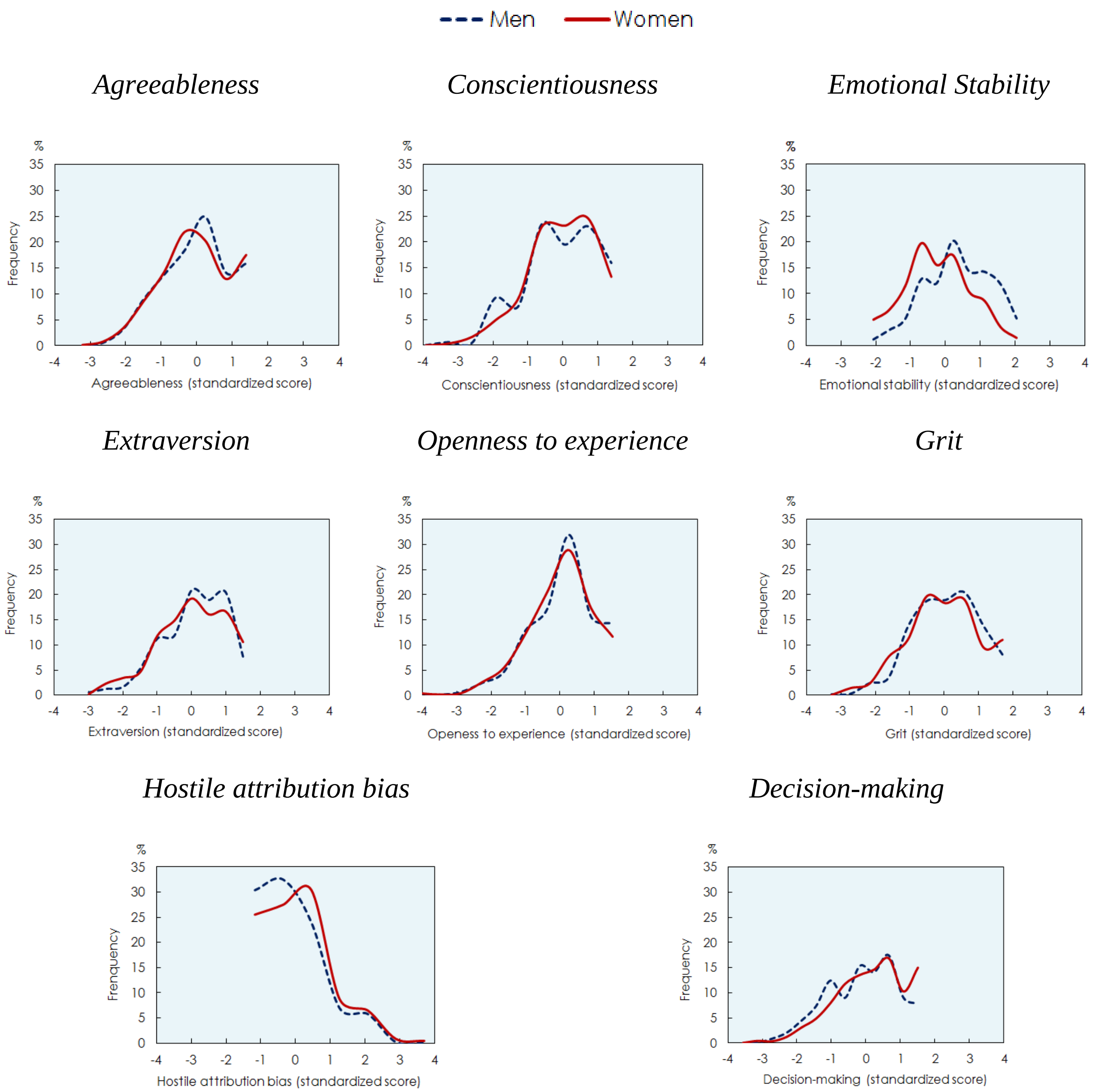


Notes: Based on Pearson's Chi-square tests, differences in the distribution of reading proficiency scores (mean of the 10 Plausible Values) are significant at the 95percent level for age and education levels (not across gender).

Source: Authors' elaboration based on Colombia STEP Household Survey.

Figure 2. Distribution of socio-emotional skills across gender

Share of individuals by socio-emotional skills scores across age groups

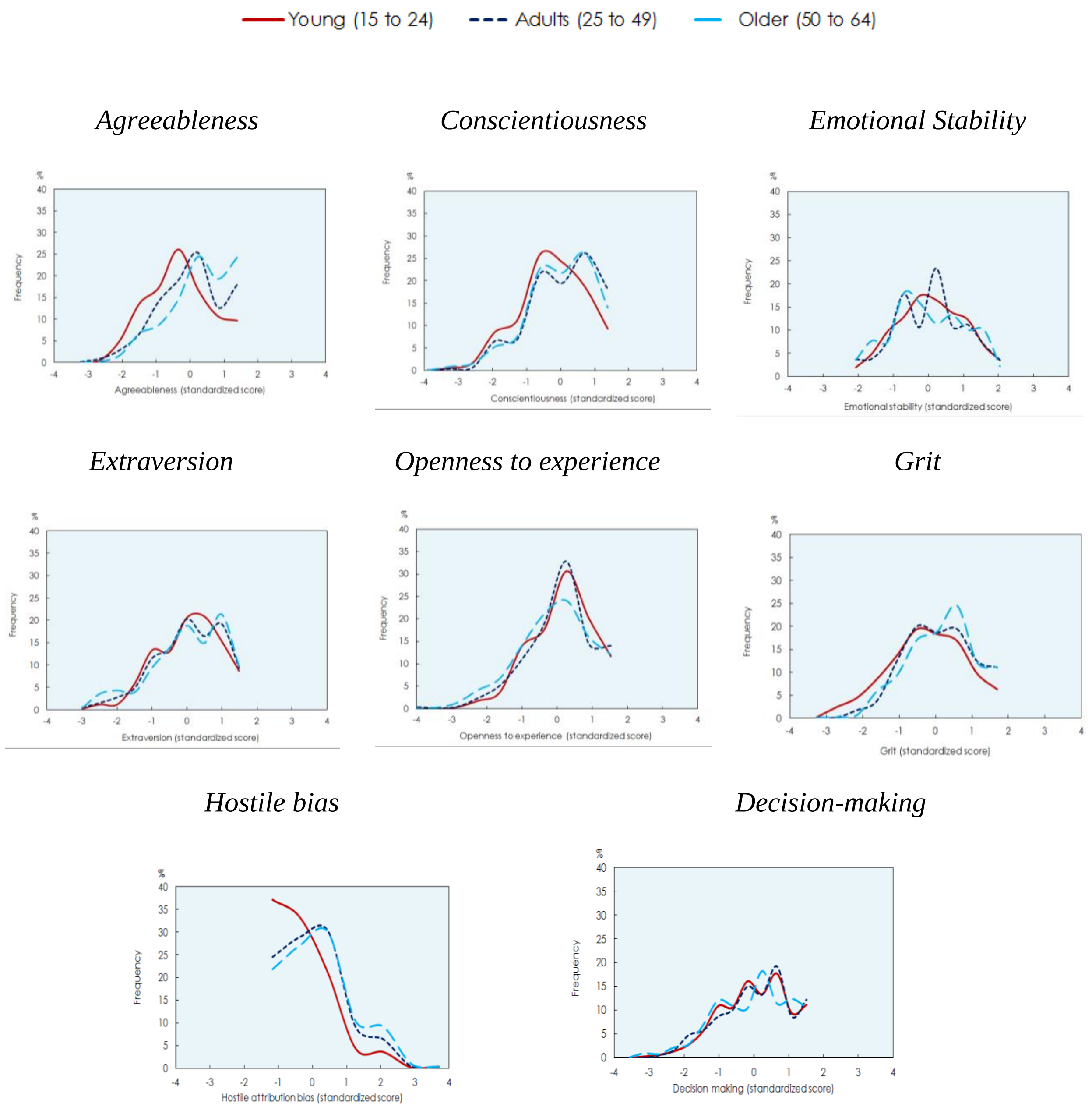


Notes: The figures above represent scatter plots with smoothed lines based on tabulations of standardized scores of socio-emotional skills and gender. Based on Pearson's Chi-square tests, differences in the distribution of conscientiousness, emotional stability, grit, decision making and hostile attribution bias are significant at the 95percent level.

Source: Authors' elaboration based on Colombia STEP Household Survey.

Figure 3. Distribution of socio-emotional skills across age groups

Share of individuals by socio-emotional skills scores across age groups

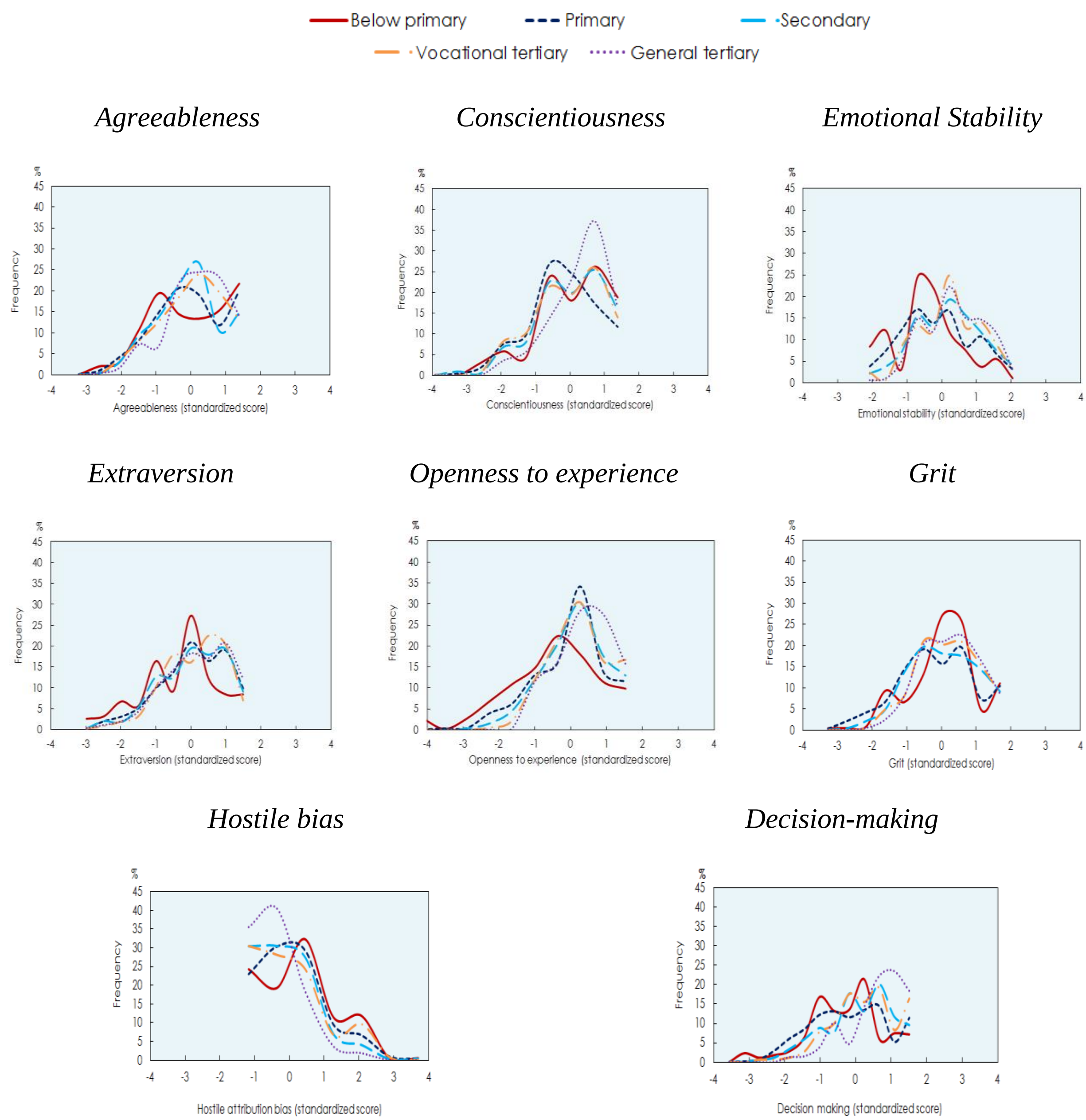


Note: The figures above represent scatter plots with smoothed lines based on tabulations of standardized scores of socio-emotional skills and age groups. Based on Pearson's Chi-square tests performed two levels by two, differences in the distribution of agreeableness, emotional stability, grit and hostile attribution bias are statistically significant at the 95percent between at least at two levels.

Source: Authors’ elaboration based on Colombia STEP Household Survey.

Figure 4. Distribution of socio-emotional skills across highest educational level completed

Share of individuals by socio-emotional skills scores across educational levels



Note: The figures above represent scatter plots with smoothed lines based on tabulations of standardized scores of socio-emotional skills and highest completed educational level. Based on Pearson's Chi-square tests performed two levels by two, differences in the distribution of are all statistically significant at the 95percent between at least at two levels.

Source: Authors' elaboration based on Colombia STEP Household Survey.

Figure 5. Structural Estimation, Variance decomposition of the scores

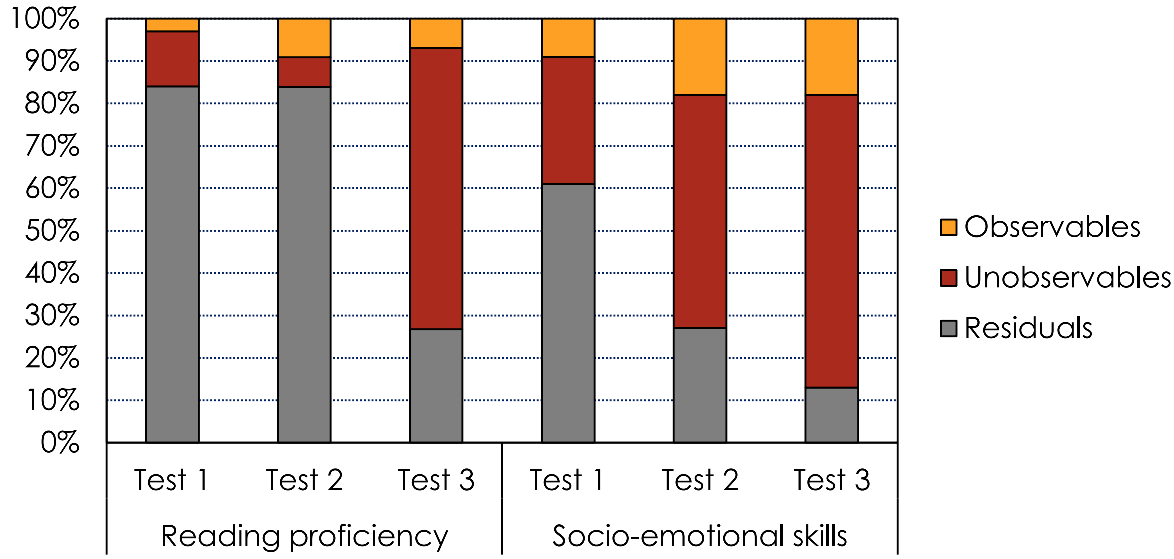


Figure 6. Probability of being active or in school

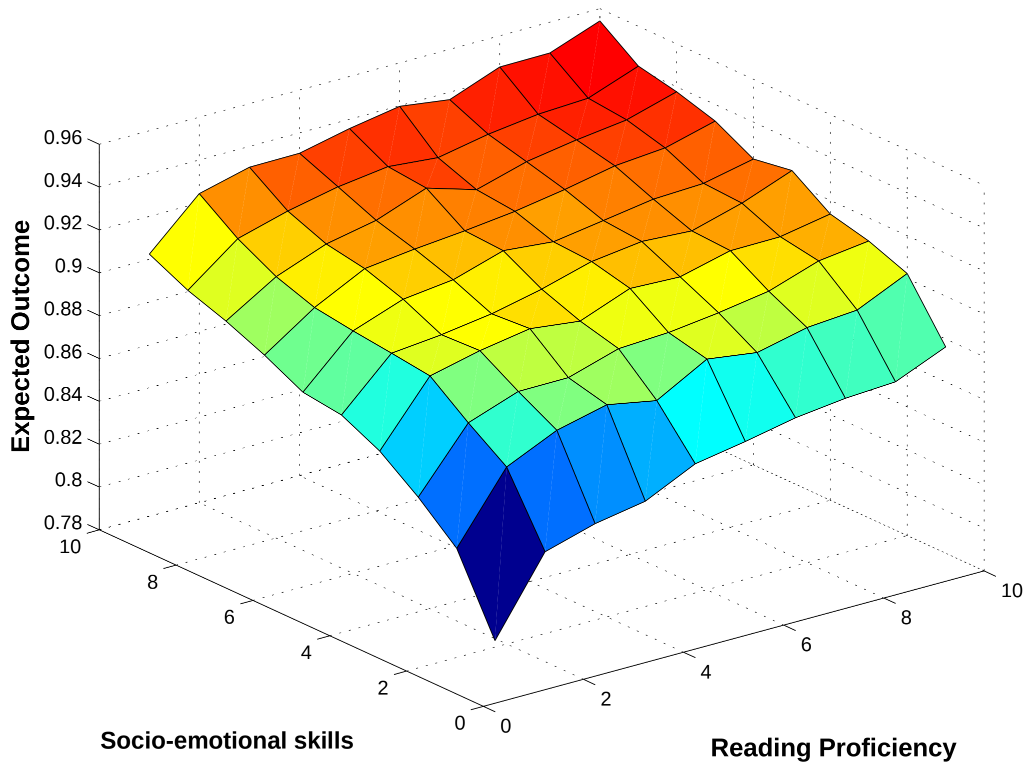


Figure 7. Probability of Having Attended Tertiary Education

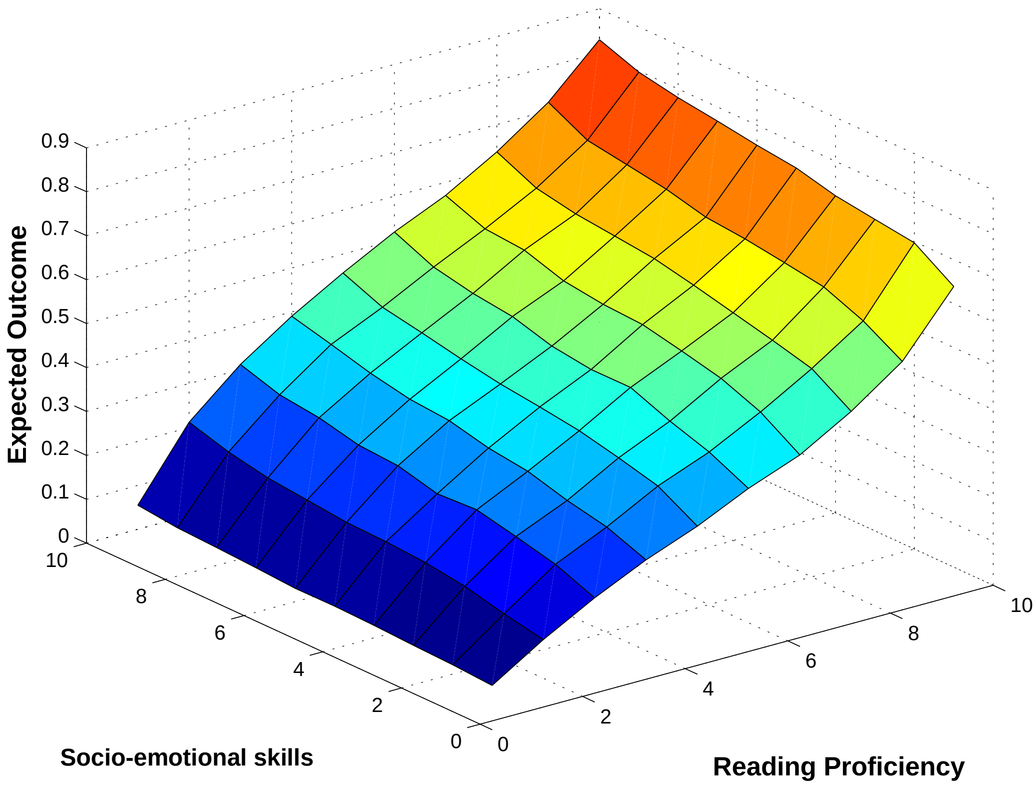


Figure 8. Probability of Being Employed

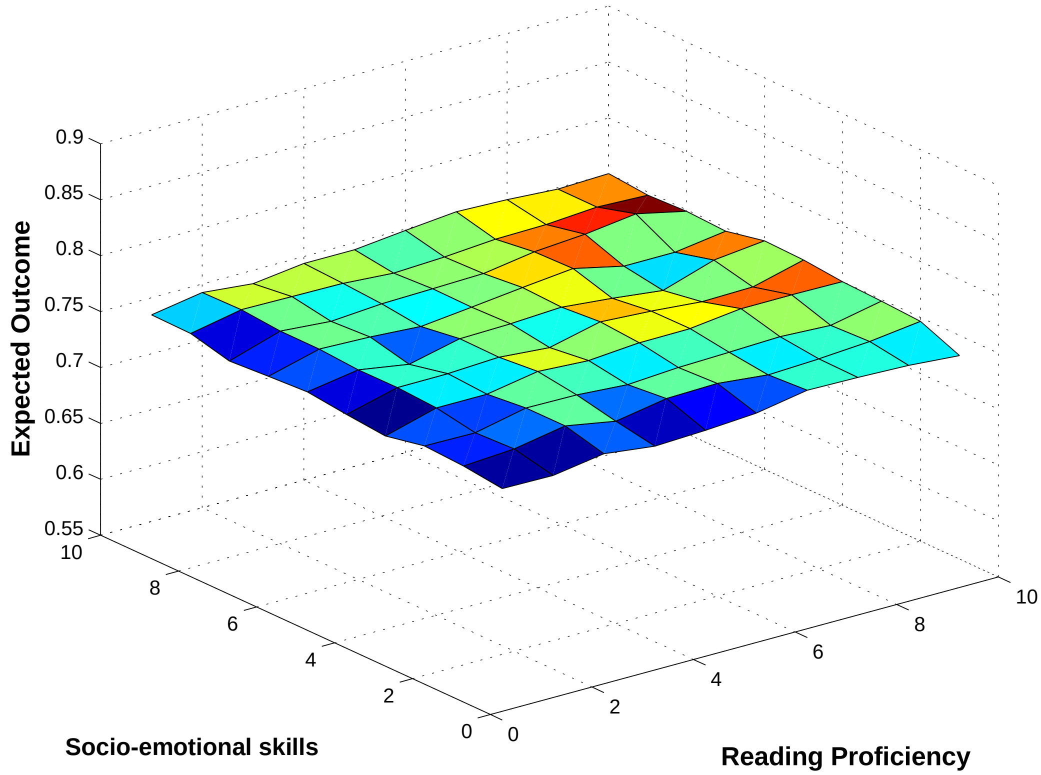


Figure 9. Probability of Having a Formal Job

Figure 10. Probability of Being a White-Collar Worker

Figure 11. Hourly Income (in CO$, Colombian pesos) from Main Job

Table 1. Partial correlations between reading proficiency and socio-emotional skills

	REA	EXT	CONS	OPE	EMO	AGR	GRI	DMG	HAB
Reading proficiency (REA)	1								
Extraversion (EXT)	0.06	1							
Conscientiousness (CONS)	0.06	0.05	1						
Openness to experience (OPE)	0.20*	0.17*	0.16*	1					
Emotional stability (EMO)	0.10*	0.10*	0.06	0.09*	1				
Agreeableness (AGR)	-0.03	0.11*	0.16*	0.20*	0.04	1			
Grit (GRI)	0.00	0.05	0.21*	0.20*	0.00	0.21*	1		
Decision making (DMG)	0.23*	0.08*	0.17*	0.29*	-0.08*	0.17*	0.21*	1	
Hostile attribution bias (HAB)	-0.17*	-0.01	-0.04	0.00	-0.17*	0.01	-0.02	-0.05	1

Note: Partial correlations with * are statistically significant at the 0.001 level.

Table 2. OLS regressions of log hourly labor earnings on cognitive skills and socio-emotional skills

Dependent variable: log hourly labor earnings

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Reading proficiency	0.178***									0.161***
	(0.05)									(0.05)
Extraversion		0.013								-0.009
		(0.04)								(0.04)
Conscientiousness			-0.023							-0.034
			(0.04)							(0.04)
Openness to experience				0.099***						0.082**
				(0.04)						(0.03)
Emotional stability					0.010					0.008
					(0.04)					(0.04)
Agreeableness						0.042				0.023
						(0.03)				(0.03)
Grit							-0.021			-0.030
							(0.04)			(0.04)
Hostile attribution bias								-0.011		-0.003
								(0.03)		(0.03)
Decision making									0.054	0.013
									(0.04)	(0.04)
Observations	1,372	1,372	1,372	1,372	1,372	1,372	1,372	1,372	1,372	1,372
R-squared	0.09	0.07	0.07	0.08	0.07	0.07	0.07	0.07	0.07	0.11

Note: Standard errors in parentheses. Significance levels are as follow: *** p<0.01, ** p<0.05, * p<0.1. All regressions are estimated using OLS and include controls for being a woman (dummy), age, age-squared, mother education (dummies, primary education is the reference category), cities of living and their metropolitan areas (dummies, Bogota-Barranquilla-Villavicencio is the reference category). The bottom and the top 1percent of the log hourly labor earnings distribution are trimmed. Measures of reading proficiency and socio-emotional skills are standardized. Regressions coefficients and standards errors of reading proficiency are the average of the 10 estimations using plausible values.

Source: Own calculations based on Colombia STEP Household Survey.

Table 3. Conditional correlations of labor market outcomes on skills, traits and schooling

Outcomes	Log hourly labor earning		Being formal worker		Being white-collar worker		Being employed		Being active or in school	Having attended tertiary education
Method	OLS		Logit		Logit		Logit		Logit	Logit
With/without schooling	Without	With	Without	With	Without	With	Without	With	Without	Without
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Reading proficiency	0.161***	0.065	0.310***	0.074	0.854***	0.442***	0.016	-0.051	0.194	1.364***
	(0.05)	(0.06)	(0.11)	(0.12)	(0.15)	(0.15)	(0.10)	(0.11)	(0.12)	(0.17)
Extraversion	-0.009	0.000	0.006	0.028	-0.026	-0.002	-0.039	-0.042	0.054	-0.058
	(0.04)	(0.04)	(0.09)	(0.09)	(0.08)	(0.09)	(0.09)	(0.09)	(0.12)	(0.09)
Conscientiousness	-0.034	-0.034	-0.014	-0.014	0.006	-0.012	0.252***	0.261***	0.254***	0.017
	(0.04)	(0.04)	(0.10)	(0.10)	(0.09)	(0.09)	(0.09)	(0.09)	(0.10)	(0.10)
Openness to experience	0.082**	0.078**	-0.100	-0.112	0.124	0.127	0.065	0.055	0.158	0.308***
	(0.03)	(0.03)	(0.09)	(0.09)	(0.09)	(0.10)	(0.09)	(0.09)	(0.10)	(0.10)
Emotional stability	0.008	-0.015	0.132	0.084	0.033	-0.051	0.103	0.085	0.093	0.329***
	(0.04)	(0.04)	(0.10)	(0.10)	(0.09)	(0.10)	(0.09)	(0.09)	(0.12)	(0.10)
Agreeableness	0.023	0.015	-0.056	-0.070	-0.042	-0.019	-0.091	-0.094	-0.178	0.010
	(0.03)	(0.03)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)	(0.11)	(0.10)
Grit	-0.030	-0.043	-0.096	-0.129	0.076	0.052	0.026	0.024	0.017	0.024
	(0.04)	(0.04)	(0.09)	(0.10)	(0.09)	(0.10)	(0.09)	(0.09)	(0.11)	(0.10)
Hostile attribution bias	-0.003	0.023	-0.203**	-0.157*	-0.114	-0.023	-0.057	-0.046	-0.098	-0.333***
	(0.03)	(0.03)	(0.09)	(0.09)	(0.09)	(0.10)	(0.08)	(0.08)	(0.09)	(0.10)
Decision making	0.013	-0.007	0.060	0.013	0.112	-0.017	-0.178*	-0.191**	-0.034	0.374***
	(0.04)	(0.04)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)	(0.11)	(0.09)
Education: below primary		0.015		0.533		0.667		-0.366		
		(0.11)		(0.43)		(0.53)		(0.34)		
Education: upper secondary		0.289***		1.021***		1.191***		0.019		
		(0.10)		(0.24)		(0.31)		(0.23)		
Education: vocational tertiary		0.371***		1.356***		2.030***		0.304		
		(0.11)		(0.27)		(0.31)		(0.27)		
Education: general tertiary		0.880***		1.796***		4.132***		0.380		
		(0.15)		(0.35)		(0.38)		(0.39)		
Observations	1,372	1,372	1,577	1,576	1,801	1,801	2,117	2,117	2,356	1,717

Note: Standard errors in parentheses. Significance levels are as follow: *** p<0.01, ** p<0.05, * p<0.1. Conditional correlations are computed from Ordinary Least Square Regressions (OLS) for labor earnings and Logit regressions for labor supply and education outcomes. The bottom and the top 1percent of the log hourly labor earnings distribution are trimmed. OLS calculations control for, being a woman (dummy), age, age-squared, mother education (dummies, primary education is the reference category), cities of living and their metropolitan areas (dummies, Bogota-Barranquilla-Villavicencio is the reference category). Logit regressions control for the same variables and a self-reported categorical variable on parents' involvement in ones' education at the age of 12 (3 levels). Odd ratios are reported for logit regressions. Measures of reading proficiency and socio-emotional skills are standardized. Regressions coefficients and standards errors of reading proficiency are the average of the 10 estimations using plausible values.
Source: Own calculations based on Colombia STEP Household Survey.

Table 4. Conditional correlations of labor market outcomes on skills and socio-emotional skills, across subsamples

Outcomes	Log hourly labor earning						Being active or in school					
Method	OLS						Logit					
Subsamples	Men	Women	Younge r	Older	Less educated	More educated	Men	Women	Younger	Older	Less educated	More educated
Reading proficiency	0.160**	0.140**	0.134*	0.173***	0.019	0.180**	-0.013	0.271**	0.449**	0.095	0.282**	0.208
	(0.07)	(0.06)	(0.07)	(0.06)	(0.05)	(0.08)	(0.25)	(0.13)	(0.18)	(0.15)	(0.14)	(0.20)
Extraversion	-0.030	0.024	0.011	-0.011	0.007	0.008	-0.039	0.121	0.356**	-0.125	0.235*	0.041
	(0.05)	(0.05)	(0.05)	(0.04)	(0.04)	(0.05)	(0.22)	(0.12)	(0.14)	(0.14)	(0.14)	(0.14)
Conscientiousness	-0.037	-0.017	-0.102*	0.044	-0.040	-0.004	0.065	0.299***	0.310**	0.206*	0.375***	0.048
	(0.05)	(0.05)	(0.05)	(0.04)	(0.04)	(0.05)	(0.17)	(0.11)	(0.14)	(0.12)	(0.14)	(0.13)
Openness to experience	0.130***	0.046	0.030	0.152***	0.065	0.066*	0.092	0.200*	0.339**	0.207*	0.080	0.212
	(0.04)	(0.05)	(0.04)	(0.05)	(0.04)	(0.04)	(0.21)	(0.12)	(0.15)	(0.12)	(0.14)	(0.13)
Emotional stability	-0.031	0.047	-0.013	0.036	0.011	-0.050	-0.163	0.149	0.114	0.133	0.057	0.149
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.28)	(0.12)	(0.13)	(0.15)	(0.15)	(0.15)
Agreeableness	0.028	0.013	-0.006	0.037	-0.003	0.050	-0.229	-0.185*	-0.153	-0.223*	-0.363***	-0.024
	(0.04)	(0.05)	(0.04)	(0.04)	(0.04)	(0.04)	(0.26)	(0.11)	(0.15)	(0.13)	(0.13)	(0.16)
Grit	-0.033	-0.040	-0.040	-0.059	-0.015	-0.084	0.131	0.022	-0.049	0.090	0.032	0.086
	(0.06)	(0.06)	(0.06)	(0.05)	(0.05)	(0.05)	(0.24)	(0.12)	(0.13)	(0.15)	(0.14)	(0.13)
Hostile attribution bias	-0.044	0.029	0.048	-0.053	0.036	-0.008	-0.182	-0.083	-0.095	-0.068	0.145	-0.239**
	(0.04)	(0.05)	(0.05)	(0.04)	(0.05)	(0.05)	(0.21)	(0.10)	(0.14)	(0.12)	(0.14)	(0.11)
Decision making	0.046	-0.030	0.033	0.019	-0.024	0.006	-0.266	-0.000	-0.082	0.019	-0.001	-0.100
	(0.05)	(0.05)	(0.05)	(0.04)	(0.04)	(0.05)	(0.27)	(0.12)	(0.16)	(0.13)	(0.15)	(0.14)
Observations	686	686	678	694	438	934	933	1,369	1,233	1,123	864	1,492
R-squared	0.12	0.09	0.13	0.16	0.06	0.14						

Note: Standard errors in parentheses. Significance levels are as follow: *** p<0.01, ** p<0.05, * p<0.1. Younger and older are individuals respectively under 35 and 35 and above. Less educated individuals have respectively complete primary education or below and more educated ones have at least completed upper secondary education. The bottom and the top 1percent of the log hourly labor earnings distribution are trimmed. OLS calculations control for, being a woman (dummy), age, age-squared, mother education (dummies, primary education is the reference category), cities of living and their metropolitan areas (dummies, Bogota-Barranquilla-Villavicencio is the reference category). Logit regressions control for the same variables and a self-reported categorical variable on parents' involvement in ones' education at the age of 12 (3 levels). Odd ratios are reported for logit regressions. Measures of reading proficiency and socio-emotional skills are standardized. Regressions coefficients and standards errors of reading proficiency are the average of the 10 estimations using plausible values.

Table 5. IV estimates (second stage) of labor market outcomes on skills and traits

Instrumented skills/traits	(1) Log hourly labor earning	(2) Being formal worker	(3) Being white-collar worker	(4) Being employed	(5) Being active or in school	(6) Having attented tertiary education
Reading proficiency (posterior means of PV)	0.483**	0.119	0.196**	-0.038	0.042	0.612***
	(0.22)	(0.10)	(0.08)	(0.08)	(0.07)	(0.10)
Extraversion	-1.492	-2.384	4.461	-0.061	0.420	1.316
	(2.09)	(10.40)	(36.87)	(0.51)	(0.85)	(1.06)
Conscientiousness	-0.429	-0.143	-0.283**	0.161	0.070	-0.467**
	(0.31)	(0.11)	(0.12)	(0.12)	(0.09)	(0.21)
Openness to experience	-0.020	0.297	-0.564	-0.736	-0.307	1.708
	(0.74)	(0.62)	(1.10)	(1.09)	(0.40)	(1.70)
Emotional stability	0.170	-0.018	0.125	0.104	0.250	0.365*
	(0.34)	(0.16)	(0.17)	(0.25)	(0.25)	(0.19)
Agreeableness	-0.051	0.070	-0.984	0.243	-0.125	-1.025
	(1.19)	(0.67)	(1.37)	(0.33)	(0.25)	(0.85)
Grit	0.419	0.155	0.332**	-0.093	0.022	0.539***
	(0.31)	(0.17)	(0.16)	(0.14)	(0.08)	(0.17)
Hostile attribution bias	-0.554	-0.205	-0.458**	0.112	-0.001	-0.659***
	(0.41)	(0.17)	(0.21)	(0.13)	(0.11)	(0.24)
Decision making	-0.490	-0.052	-0.324	-0.379	-0.288	-0.009
	(0.48)	(0.20)	(0.24)	(0.46)	(0.28)	(0.35)
Observations	1,371	1,575	1,800	2,116	2,355	1,717

Note: Standard errors in parentheses. Significance levels are as follow: *** p<0.01, ** p<0.05, * p<0.1. \The bottom and the top 1percent of the log hourly labor earnings distribution are trimmed. IVs considered for reading proficiency are the age at which a person started school, and the economic situation of the household at age 12; and for each socio-emotional skill, the indicator of whether the individual lived with both parents at age 12, and the economic situation of household at age. 12. These instruments are valid in first stage per Sargan-Hansen test. Other controls are gender, age, age-squared, mother education (dummies, primary education is the reference category), cities of living and their metropolitan areas (dummies, Bogota-Barranquilla-Villavicencio is the reference category). Measures of reading proficiency and socio-emotional skills are standardized. Regressions coefficients and standards errors of reading proficiency are the average of the 10 estimations using plausible values. For presentational purposes, the coefficients from different regressions for same outcome are presented in one column.

Table 6. Socio-emotional skills Factor Estimation, Whole Sample

	Extrv-Open	Stab-Hostile	Cons-Grit-DM
Age	**0.0183***	**-0.0200****	**0.1009*****
	(0.011)	(0.009)	(0.010)
Age2	-0.0001	0.0002	**-0.0012*****
	(0.000)	(0.000)	(0.000)
Female	**-0.0826***	**-0.4869*****	**0.1085****
	(0.048)	(0.041)	(0.047)
Mother Ed. < Primary	**-0.5259*****	**-0.4857*****	**-0.3593*****
	(0.123)	(0.105)	(0.116)
Mother Ed. = Primary	**-0.2304****	**-0.2180****	**-0.1803***
	(0.099)	(0.085)	(0.093)
Mother Ed. = Secondary	-0.1561	-0.0640	-0.1184
	(0.101)	(0.086)	(0.095)
Socio-emotional Skill	**0.4412****	0.0447	1
	(0.174)	(0.047)	.
Constant	**9.2485*****	**1.7043*****	**7.6960*****
	(0.197)	(0.169)	(0.189)

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All the estimations include city dummies, coefficients not reported. Mother's education omitted category is tertiary education and beyond. N=2,372

Table 7. Reading Proficiency Factor Estimation, Whole Sample

	Language	Quant Read	PV Literacy
Age	**0.0626*** **	**0.0100**	**0.0240*** **
	(0.019)	(0.007)	(0.007)
Age2	**-0.0011*** **	**-0.0002** **	**-0.0004*** **
	(0.000)	(0.000)	(0.000)
Female	-0.0506	-0.0376	-0.0272
	(0.086)	(0.033)	(0.031)
Mother Ed. < Primary	**-1.1520*** **	**-1.0506*** **	**-0.9422*** **
	(0.216)	(0.080)	(0.075)
Mother Ed. = Primary	**-0.5451*** **	**-0.6397*** **	**-0.4983*** **
	(0.172)	(0.064)	(0.059)
Mother Ed. = Secondary	-0.1183	**-0.3318*** **	**-0.2433*** **
	(0.175)	(0.065)	(0.060)
Reading proficiency	**1.6001*** **	**0.9055*** **	1
	(0.057)	(0.020)	.
Constant	**25.2210*** **	**0.6072*** **	**0.3168*** **
	(0.346)	(0.129)	(0.121)

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All the estimations include city dummies, coefficients not reported. Mother's education omitted category is tertiary education and beyond. N=2,340

Table 8. Structural estimates of labor market outcomes on latent skills and traits

Outcomes	**Log hourly labor earning**	**Being formal worker**	**Being a white-collar worker**	**Being employed**	**Being active or in school**	**Having attented tertiary education**
	(1)	(2)	(3)	(4)	(5)	(6)
Reading proficiency	**0.134*** **	**0.276*** **	**0.252*** **	0.023	**0.112** **	**0.988*** **
	(0.032)	(0.052)	(0.040)	(0.042)	(0.047)	(0.076)
Socio-emotional skill	-0.026	-0.004	0.046	0.013	**0.143*** **	**0.170*** **
	(0.028)	(0.044)	(0.035)	(0.040)	(0.045)	(0.049)
Age	**0.032*** **	**0.137*** **	**0.123*** **	**0.171*** **	**0.073*** **	0.027
	(0.012)	(0.019)	(0.013)	(0.017)	(0.016)	(0.029)
Age2	**-0.000** **	**-0.002*** **	**-0.002*** **	**-0.002*** **	**-0.001*** **	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Female	**-0.198*** **	**-0.396*** **	**0.093* **	**-0.712*** **	**-0.874*** **	-0.076
	(0.044)	(0.068)	(0.055)	(0.066)	(0.085)	(0.075)
Mother Ed. < Primary	**-0.810*** **	-0.104	**-0.939*** **	0.148	**-0.476** **	**-2.559*** **
	(0.122)	(0.180)	(0.145)	(0.167)	(0.232)	(0.267)
Mother Ed. = Primary	**-0.507*** **	-0.140	**-0.469*** **	0.033	**-0.635*** **	**-1.637*** **
	(0.103)	(0.146)	(0.113)	(0.138)	(0.209)	(0.231)
Mother Ed. = Secondary	**-0.249** **	-0.041	**-0.214* **	0.077	**-0.434** **	**-1.069*** **
	(0.106)	(0.150)	(0.114)	(0.142)	(0.212)	(0.236)
Parental Involv1		0.070	-0.058	-0.030	0.004	0.100
		(0.093)	(0.075)	(0.082)	(0.093)	(0.096)
Parental Involv2		0.035	**-0.214** **	-0.113	-0.101	-0.023
		(0.113)	(0.091)	(0.100)	(0.114)	(0.120)
Constant	**0.032*** **	**-2.223*** **	**-1.769*** **	**-2.057*** **	**1.425*** **	0.673
	(0.012)	(0.373)	(0.245)	(0.329)	(0.347)	(0.625)
Observations	1,363	1,560	2,328	2,089	2,328	1,692

Note: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All the estimations include city dummies, coefficients not reported. Mother's education omitted category is tertiary education and beyond. The estimations presented are raw coefficients.

Source: Own calculations based on the Household Survey STEP Colombia, 2012

Appendix. STEP Household Skills Measurement Survey for Colombia.

General Description

The household survey questionnaires are structured in four blocks. The first block, in line with standard household surveys, includes modules to obtain information on the household's socio-economic conditions, the demographic characteristics of household members (household roster) and household dwelling. Once the first block is administered to the household, one working age member of the household will be chosen to respond the individual-level blocks (Bocks 2-4) administered. This household member to be interviewed (called the main respondent) is selected using a standardized random approach. The main respondent is selected from among all household members aged 15 to 64 years who are present in the household or is present within three days of the first visit to the household, using a random procedure specified in the questionnaire. There is no replacement of main respondent allowed.

The survey consists of questions on labor market performance (present employment, past employment), followed by modules on workplace skills/ technical skills, and education. It also includes direct and indirect assessments to test or obtain information on different types of skills sets: a) cognitive skills (literacy tested with exercises that the respondent completes, designed by ETS); and b) socio-emotional or non-cognitive skills (such as personality traits, behaviors, time and risk preferences).

The survey covers the main urban areas of the country. The surveys comprise a sample of approximately 2,400 households that is selected in two stages. In the first stage, 200 small territorial areas (hereafter referred to as Primary Sampling Units, or PSUs) are selected as main PSU's.; in the second stage, 15 households are selected in each PSU visited. Both stages are random selections.

Content of Household Survey

<u>Household Roster</u>

(a) Dwelling and Household Characteristics
   a. Dwelling construction materials, number of rooms, source of water and energy, toilets
   b. Tenure status
   c. Inventory of household consumer goods, appliances, and vehicles, number of books
   d. Ownership of bank accounts, receipt of social benefits
   e. Household size, gender and age distribution, relationship to head, members' marital status
   f. Household members' employment status, education and literacy levels

<u>Individual Respondent</u>

(b) Education and Training
  a. Level of formal education and whether academic or vocational
  b. Field of study for highest qualification (13-15 categories)
  c. Reasons for dropping out (if applicable)
  d. Apprenticeship (y/n) and trade
  e. Number of training courses, participation in literacy courses
  f. School class rank, parental encouragement

(c) Health
  a. Overall life satisfaction
  b. Height. Weight, present or previous health problems and severity
  c. Insurance coverage

(d) Employment
  a. Employment status, whether work on own account and casual work
  b. Reason not working, job search methods, reason not looking for work (if not working)
  c. Reservation wage, occupations for which qualified (if not working)
  d. Occupation, tenure, industry, hours worked, other occupations for which qualified
  e. Class of worker (wage/salary, daily or piecework, self-employed with(out) employees)
  f. Wage, salary, or profits per time period, in-kind payments
  g. Employer (government, individual, domestic or foreign firm, NGO)
  h. Establishment size, social benefit coverage

(e) Job Skill Requirements
  a. Inventory of reading tasks performed on job (or in general), length of longest document read
  b. Inventory of writing tasks performed on job (or in general), length of longest written document
  c. Inventory of math tasks performed on job (or in general)
  d. Whether lack of reading and writing skills hindered employment, promotion, or pay raise
  e. Frequency of difficult problem solving on job
  f. Level of involvement with customers, clients, students, or public on job
  g. Make formal presentations as part of job
  h. Supervisory responsibilities, job autonomy, repetitiveness, continuous learning
  i. Level of physical job demands
  j. Inventory of technology use on job (including computer use and inventory of software use)
  k. Computer use outside work and inventory of software use
  l. Whether lack of computer skills has hindered employment, promotion, or pay raise
  m. Usefulness of own studies at school for current job
  n. Level of education and related job experience required for job, length of job learning time
  o. Job search skills, whether employer required formal credential or other proof of skills

(f) Personality, Behavior and Preferences
    a. Thirty-one personality items on the frequency of diagnostic behaviors (e.g., extroversion)
    b. Seven-item risk preference scale
(g) Language and Family Background
    a. Native language, other specific language proficiency
    b. Mother's and father's educational attainment
    c. Family size, composition, and socio-economic status when 12 years old, adverse family events
    d. Experience as child laborer, occupation
(h) Reading Literacy Test
    a. Core
    b. Reading Components
    c. Exercise booklets
(i) Interviewer Impressions
    a. Comprehension of questions, reliability and candor, distractions