Beyond Qualifications
Returns to Cognitive and Socioemotional Skills in Colombia*

Pablo Acosta, World Bank and IZA
Noël Muller, World Bank
Miguel Sarzosa, Purdue University

June 2017

Abstract

A growing literature has established the importance of both cognitive skills (mental abilities) and socioemotional skills (personality traits and behaviors) in shaping labor market outcomes, beyond education levels. However, this literature has not reached yet a consensus on the respective role of this skills sets. This paper aims to fill this gap by examining how adults' cognitive and socioemotional skills relate to a range of labor market outcomes in Colombia. Using a 2012 household skills survey, it finds that cognitive and socioemotional skills correlate more strongly with distinct outcomes: socioemotional skills appear to play a strong role in labor market participation and employment while cognitive skills are strongly related with higher earnings, holding a formal job, and working in occupation that requires high qualifications. Both types of skills are strongly related with tertiary education. The analysis applies standard econometric techniques as a benchmark and structural estimations to correct for the measurement error of skill constructs.

JEL codes: J24, J31, I24.

Keywords: Colombia, cognitive skills, socioemotional skills, latent skills, unobserved heterogeneity, labor market outcomes.

* The authors are grateful to Sergio Urzua, Margaret Grosh, Wendy Cunningham, Alexandria Valerio, Maria Laura Sanchez Puerta, Mónica Parra Torrado, Juan D. Barón, and Ana Maria Oviedo for their comments and frequent interactions with the authors while developing this paper. Correspondence concerning this article should be addressed to: Pablo Acosta, pacosta@worldbank.org; Noël Muller, nmuller@worldbank.org; and Miguel Sarzosa, msarzosa@purdue.edu.

1. Introduction

A large body of evidence has documented that direct measures of skills provide more adequate estimations of individuals' differences in potential productive capacity than the quantity of education they receive (Hanushek 2015). People are distinguished by the broad sets of abilities, accumulated over their life course, that shape their decisions and success in the labor market (Almlund *et al.* 2011; Borghans *et al.* 2008). These broad sets of skills fall into two overlapping categories: (1) cognitive skills[1]—mental abilities such as comprehension or reasoning—and (2) socioemotional skills—personality traits, behaviors, attitudes, and beliefs.[2] Common proxies such as years of education do not capture well cross-country differences in skills acquisition at school and the development of these skills outside the classroom (Hanushek and Woessman 2008). In fact, employer surveys from Latin America and beyond confirm that socioemotional and advanced cognitive skills are valued, as much as, or even more than, qualifications alone (Cunningham and Villaseñor 2016; Bassi *et al.* 2012).

This paper examines the impact of individuals' skills on labor outcomes in the context of Colombia, an upper-middle-income country. Using a 2012 cross-sectional household survey providing direct measures of skills and rich background information, we investigate (1) the current levels and distribution of cognitive and socioemotional skills in Colombia's urban working-age population, and (2) the degree to which certain types of cognitive and socioemotional skills relate to a range of labor market outcomes and educational trajectories. Our findings complement the existing evidence that comes mainly from high-income countries (mostly Western Europe and the United States) and a fistful of developing countries.

We use three methodologies to compare results. We use first conventional methodologies (OLS for labor earnings, Logit for binary outcomes) that the literature uses. Second, we use a structural model that identifies latent cognitive and socioemotional skills as sources of unobserved heterogeneity and correct measurement errors. This acknowledges the fact that the scores obtained from surveys are manifest variables from the true latent measures of skills (Bartholomew *et al.*, 2011).[3] Survey measures of skills-based, self-reported items or tests have proven to be fallible in capturing individuals' true, or latent, skills (Heckman and Kautz 2012;

---

[1] Cognitive skills are understood here as intelligence or mental abilities. Two levels can be distinguished: (1) basic or foundational cognitive skills includes academic knowledge such as literacy or numeracy, and (2) higher-order (advanced) cognitive skills involve more complex thinking such as critical thinking or problem solving (Neisser et al. 1996; Cattell 1987)

[2] The term socioemotional skills are understood as the set of abilities that enable individuals to navigate personal and social situations effectively (Guerra, Modecki, and Cunningham 2014). Economists commonly considered behavioral characteristics and personality traits under the umbrella of "non-cognitive skills", "behavioral skills", or "soft skills" or leave the distinction unexplained. Socioemotional skills encompass behaviors and attitudes that are consistent patterns of thoughts, feeling, and conduct (such as commitment, discipline, or the ability to work in a team) while personality traits (such as self-confidence, perseverance, and emotional stability) are broad facets relatively stable over time that influence behaviors and attitudes (Borghans et al 2008; Almlund et al. 2011). For the sake of simplicity, we classify here behaviors and personality traits under the single rubric of socioemotional skills.

[3] Structural estimations of latent skills are used by Heckman, Stixrud, and Urzúa (2006) and Urzúa (2008), among others.

Almlund *et al.* 2011; Borghans *et al.* 2011). We complement these results with IV estimations to correct measurement error, omitted variable bias, and explore causality.

We find that both cognitive and socioemotional skills matter for favorable labor market outcomes in the Colombian context, although they have distinct roles. Cognitive skills are greatly associated with higher earnings and holding a formal job or a high-qualified occupation; for example, a switch from the first to the tenth decile in cognitive skills correlates with an increase of 28 percentage points in the probability of holding a formal job (i.e. more than the double). By contrast, socioemotional skills appear to have modest influence on these outcomes but play a stronger role in labor market participation and employment: an adult switching from the first to the tenth decile is 9 percent more likely of working, looking for a job, or studying (5.8 percentage points for cognitive skills). Both types of skills, especially cognitive, are largely associated with the attainment of tertiary education: adults in the lowest decile of both skills are virtually unlikely of having gone to college (only 1.5%) while adults in the highest levels of both type of skills have 83 percent of probability of having done so. These findings roughly hold true across types of estimates for approaches using both disaggregated measures of skills.

These findings corroborate evidence from high-income countries, which offer several explaining channels. The relationship between cognitive skills and labor earnings and quality jobs is well established. Both in high-income countries and lower-income countries, cognitive skills reflect how people think and solve problem, which make them more productive and in turn are rewarded by higher returns (Murnane, Willett, and Levy 1995; Glewwe 2002, Hanushek and Woessmann 2008). People acquire most of their cognitive skills at school but these skills have an impact on outcomes beyond education levels. In Colombia we find that raising reading proficiency by a standard deviation is correlated with an increase of 15 percent in hourly labor earnings from the main job (8 percent for openness to experience, classified as socioemotional skills). This consistent with estimates for the United States that range from 10 to 20 percent increase of labor earnings (Hanushek 2015).

Research on Germany and the United States also shows that socioemotional skills influence preferences and the propensity to be motivated, be disciplined, and adopt an optimistic attitude while searching for a job or deciding to participate to the labor market (Mohanty 2010; Wichert and Pohlmeier 2010; Uysal and Pohlmeier 2011; Caliendo, Cobb-Clark, and Uhlendorff 2015). Socioemotional skills may also enhance the probability of receiving and accepting a job offer and of keeping a job (Mueller and Plug 2006).

The remainder of this paper is organized as follows. Section 2 provides definitions of cognitive and socioemotional skills. Section 3 reviews the empirical literature on the relationship between these skills and labor market outcomes; section 4 describes the data; and section 5 introduces the empirical strategy. The results are presented in section 6 for descriptive statistics on the levels and distributions of skills and in section 7 for ordinary least square (OLS), instrument variables (IV) estimates, and the structural estimations. The final section offers our conclusions.

2. Literature Review

2.a The Role of Skills and Traits in Labor Earnings

According to studies conducted since the mid-1990s, both cognitive skills and personality traits affect the labor earnings of the overall population, although with relatively larger effects for cognitive skills.

In the United States, cognitive abilities have long been the dominant factor determining labor earnings. In a large number of studies, higher levels of cognitive skills measured by intelligence quotient (IQ) or standardized tests of basic cognitive skills such as mathematics, reading, and vocabulary predicted higher wages, even when taking into account other factors that might also influence earnings.[4] Similar results are found in other high-income countries such as the United Kingdom (McIntosh and Vignoles 2001), Canada (Finnie and Meng 2001), and in more than 20 other member countries of the Organisation for Economic Co-operation and Development, or OECD (Hanushek *et al.* 2015).

In light of findings from program experiments and employer surveys, studies have begun to account for measures of socioemotional abilities and personality traits, in addition to cognition ones, in order to investigate their influence on labor earnings.[5] This burgeoning literature reveals that socioemotional abilities are at least as important as cognitive skills in determining labor earnings in many high-income countries such as the United States, Germany, the Netherlands, and Sweden.[6] Among the "Big Five traits"[7] used in the majority of empirical studies, conscientiousness and traits related to emotional stability (locus of control and self-esteem) are the most associated with job performance and wages in the United States and Western European countries (Barrick and Mount 1991; Bowles, Gintis, and Osborne 2001a, Heckman, Stixrud, and Urzúa 2006).[8] Using measures of socioemotional skills based on school evaluations, Carneiro,

---

[4] See Herrnstein and Murray 1994; Murnane, Willett, and Levy 1995; Gottfredson 1997; Mulligan 1999; Murnane *et al.* 2000; Altonji and Pierret 2001; Cawley, Heckman, and Vytlacil 2001; Lazear 2003; Hanushek and Woessmann 2008.

[5] For example, studies of General Educational Development (GED) recipients in the United States (high school dropouts who, by passing the GED exams, are certified as having a high school equivalent education) have served as an ideal natural experiment to confirm the crucial role of s*ocio-emotional* skills. GED recipients show higher basic cognitive skills than non-GED high school dropouts but earn, on average, the same wages. The poor labor market performances of GED beneficiaries are interpreted to originate from lower levels of s*ocio-emotional* skills, which are valued by the labor market. Being a GED graduate is a mixed signal that characterizes its recipients as smart but unreliable (Heckman and Rubinstein 2001).

[6] The following is a non-exhaustive list of sources claiming this result in the United States and Western Europe: Bowles, Gintis, and Osborne (2001b); Nyhus and Pons (2005); Osborne-Groves (2005); Heckman, Stixrud, and Urzúa (2006); Mueller and Plug (2006); Borghans, ter Weel, and Weinberg (2008); Heineck and Anger (2010); Lindqvist and Vestman (2011); and Segal (2013).

[7] A widely used taxonomy of broad families of personality traits, which include the following dimensions: openness to experience, conscientiousness, extraversion, agreeableness, and emotional stability (John and Srivastava 1999)

[8] The Big Five personality traits are openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism (opposite of emotional stability. This taxonomy summarizes a large number of distinct, more specific personality traits, behaviors, and beliefs (Goldberg 1993). 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)

Crawford, and Goodman (2007) and Segal (2013) find positive and significant associations between behaviors in childhood and adult wages in the United Kingdom and the United States. The same relationship was found between leadership abilities in youth and adult wages in the United States and Sweden (Kuhn and Weinberger 2005; Lindqvist and Vestman 2011).

Evidence suggests that skills that could be put at use across a broader array of occupations are more greatly rewarded. For example, personality traits such as conscientiousness and grit seem to matter for a wide spectrum of job complexity (Barrick and Mount 1991; Duckworth *et al.* 2007). And yet more complex jobs—that is, more demanding in information processing such as scientists and senior managers—require high-order cognitive skills that could not be used in other occupations (Schmidt and Hunter 2004). Using data for siblings in the United States, Fletcher (2013) found that extraversion shows a large and robust association with earnings that could reflect the recent change in the composition of occupations in the United States—namely, the increase in service jobs and the requirement for social interactions in the workplace. This finding is in line with that of Borghans, ter Weel, and Weinberg (2013) who document the importance of people skills in the labor markets of the United States, Germany, and the United Kingdom.[9] Higher levels of socioemotional abilities appear even more important for occupations requiring low-order cognitive skills, especially in the services sector (Bowles, Gintis, and Osborne 2001b).

The returns to higher levels of cognitive and socioemotional abilities differ across population subgroups and job types. There are often sizable differences across gender in the key personality traits with the highest rewards, although it is difficult to draw common patterns over studies.[10] For example, Osborne-Groves (2005) finds that locus of control, aggression, and withdrawal are strong predictors of wages for white women in the United States and the United Kingdom. Mueller and Plug (2006) find that agreeableness and conscientiousness seem to be more rewarding for women in the United States, while Heineck and Anger (2010) find that extraversion and agreeableness negatively affect women's wages in Germany.

The returns to skills differ across type of work as well—namely, between salaried workers and the self-employed. Individuals with high-order cognitive skills (learning aptitudes and success as a salaried worker), a tendency to break the rules, and high self-esteem in adolescence are more likely to become successful long-term entrepreneurs in the United States (Levine and Rubinstein 2013). In the Netherlands, language and clerical abilities have a stronger impact on employees' wages, whereas mathematical ability, technical ability, and extraversion in early childhood are

as opposed to the extent that the environment (i.e., chance, fate, luck) controls their lives (external control)" (Heckman, Stixrud, and Urzúa 2006).

[9] People skills are defined as "the ability to effectively interact with or handle interactions with people, ranging from communication with to caring for to motivating them" (Borghans, ter Weel, and Weinberg 2013).

[10] Heckman, Stixrud, and Urzúa (2006) find only slight variations in the effect of locus of control and self-esteem on earnings. Differences 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).

more valuable for entrepreneurs (Hartog, van Praag, and van der Sluis 2010). Moreover, entrepreneurs with a balance in abilities across different fields—that is, a jack of all trades—have higher incomes vis-à-vis salaried workers (Lazear 2005; Hartog, van Praag, and van der Sluis 2010).

Evidence for Latin American countries is scant in this topic. Pioneer work by Psacharopoulos and Velez (1992) showed that reasoning abilities and cognitive achievement (general knowledge) was strongly associated with earnings of Colombian workers in Bogota in 1988, mostly through higher educational attainment. Bassi *et al.* (2012), based on cross-sectional data for young adults in their late 20s in Argentina and Chile, found that self-efficacy is the ability that predominates the association with higher wages in both countries, with stronger effects for workers with postsecondary degrees. Díaz, Arias, and Tudela (2012) found that factors of basic cognitive skills—capturing language abilities and mathematical problem solving—and personality traits are equality valued in the Peruvian labor market for the working-age population. Specifically, grit (perseverance and passion for reaching long-term goals) and emotional stability have a high positive influence on earnings, while agreeableness shows a negative association. Although employers report valuing interpersonal skills such as teamwork, the Peruvian urban labor market does not seem to reward cooperation.

2.b The Role of Skills and Traits in Labor Supply Outcomes

Skills, especially socioemotional ones, influence individuals' participation in the labor market and probability of holding a job. As on earnings, conscientiousness, extraversion and locus of control have a large positive effect on labor participation in the United States and Germany (Barrick and Mount 1991; Gallo *et al.* 2003; Caliendo, Cobb-Clark, and Uhlendorff 2010; Wichert and Pohlmeier 2010). By contrast, neuroticism and openness to experience have a negative effect in Germany, whereas agreeableness has a negative effect only on the labor force participation decisions of married women and no effect on other population subgroups (Wichert and Pohlmeier 2010). In the United States, a man who moves from the 25th to the 75th percentile of the distribution of locus of control and self-esteem would increase his probability of being employed at age 30 by 15 percent (Heckman, Stixrud, and Urzúa 2006). Behaviors of children in the United Kingdom affect significantly the probability of having work as an adult. Although hostility toward adults in childhood has a negative impact on the probability of being in employed in one's adult years, anxiety toward acceptance by adults has a positive and significant impact on employment status (Carneiro, Crawford, and Goodman 2007). A potential explanation is that children who are maladjusted on this dimension are judged by their teachers to be overzealous, which may be better rewarded in the labor market. In Sweden, men with a lower level of leadership skills have a higher probability of being unemployed than men with lower low-order cognitive abilities (Lindqvist and Vestman 2011).

Personality traits also drive occupational choices. Individuals partly select occupations that correspond to their orientations such as being a caring or a direct person in adolescence

(Borghans, ter Weel, and Weinberg 2008, 2013). Cobb-Clark and Tan (2011) find that personality traits have a substantial effect on the probability of employment in many occupations, with gender specificities. The combination of skills and traits rather than single attributes also determines occupational outcomes. Kern *et al.* (2013) found that disagreeable intelligent individuals achieved higher occupational status, whereas disagreeable low-intelligent men were more likely to be unemployed or to work at a lower-status job.

2.c The Role of Skills in Schooling Decisions

Measures of cognitive and socioemotional skills influence schooling decisions and a range of educational outcomes (Almlund *et al.* 2011; OECD, 2015). Cunha, Heckman, and Schennach (2012) estimate that 12 percent of the variance in educational attainment is explained by personality measures, and 16 percent is accounted for by cognitive ability measures. Using longitudinal surveys of children in the United Kingdom, the United States, and Canada, Duncan *et al.* (2007) discovered that mathematics, reading, and attention skills were strong predictors of later academic achievements. By contrast, their measures of socioemotional skills at school entry had limited power in explaining educational success.[11]

Recent literature has shown that some personality traits like conscientiousness, self-discipline and grit are better predictors of the academic performance in the United States than IQ (Duckworth and Seligman 2005; Duckworth *et al.* 2007; Almlund *et al.* 2011). Openness to experience also affects educational attainment. It predicts school attendance and the level of difficulty of the courses selected. On the other hand, emotional stability—as captured by self-esteem and locus of control—influences the likelihood of graduating from high school and from a four-year college (Heckman, Stixrud, and Urzúa 2006).

Technical abilities, a subset of cognitive skills, influence the probability of going to college. By contrast to cognitive and socioemotional skill levels, in the United States a higher level of vocational ability is associated with a lower probability of attending a four-year college because individuals with higher technical skills expect higher returns from a vocational education (Prada 2013; Prada and Urzúa 2014).

Finally, DiPrete and Jennings (2012) provide evidence of the existence of 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 socioemotional skills, but girls begin school with more advanced social and behavioral skills and their skill advantage grows over time.

3. Data

[11] This could be explained by the fact that those measures of s*ocioemotional* skills influence measures of cognitive skills and therefore underestimate their effect.

The analysis in this paper is based on the Skills Toward Employment and Productivity (STEP) Household Survey, a multicountry study led by the World Bank. The survey covers a wide range of background information, similar to a standard household survey, which includes demographics, education, employment and compensation, household wealth, and household size and composition (World Bank 2014). In addition, a randomly selected individual in each household between the ages of 15 and 64 is further surveyed and tested on information related to basic cognitive skills, socioemotional skills, personal health, and use of skills on and off the job.

The STEP Household Survey of Colombia is representative of the country's 13 main cities and their metropolitan areas—that is, it covers the large majority of Colombia's urban population and is the area widely used by labor market household surveys in the country. The sample size is 2,617. The distribution in age, gender, and education attainment is similar to that for national household surveys for the same urban areas.

Measures of cognitive skills. Unless otherwise stated, the following sections of this paper use a measure associated with reading proficiency, a basic cognitive skill, produced from an advanced test developed by the Educational Testing Service (ETS).[12] *Reading proficiency* is defined as the ability to "understand, evaluate, use and engage with written texts to participate in society, to achieve one's goals, and to develop one's knowledge and potential" (OECD 2012). In this respect, reading proficiency is a broader construct than "reading," narrowly understood to be a set of strategies for decoding written text. It is intended to encompass the range of cognitive strategies (including decoding) that adults must bring into play to respond appropriately to a variety of texts of different formats and types in the range of situations or contexts in which they read (OECD 2013). To increase the accuracy of the cognitive measurement, the survey provides a set of 10 "plausible values" that are unbiased estimations of the plausible range of reading proficiency for groups of individuals (Von Davier, Gonzalez, and Mislevy 2009; OECD 2013; ETS 2014).[13]

Measures of socioemotional skills. The survey provides six measures of personality traits (relatively enduring patterns of thinking, feeling, and conduct) and two measures of behaviors and attitudes (how individuals manage interpersonal and social situations). The core of the socioemotional skills inventory is based on the Big Five model. The inventory also includes items that allows measurement of grit (i.e., a trait of perseverance and motivation to meet long-term goals (Duckworth *et al.* 2007)), hostile attribution bias (i.e., a tendency to interpret others' intents as hostile, which in return fosters one's antisocial and aggressive behavior (Dodge 2003)) and the Melbourne Decision Making Scale, which captures coping strategies for decisional conflict

[12] The reading proficiency test is comparable to the one produced by the Program for the International Assessment of Adult Competencies (PIAAC), another large-scale survey covering 24 OECD countries.
[13] Plausible values are multiple imputations, drawn after data collection, by combining test results with all available background information such as gender, age, and education. This procedure, based on item response theory, allows one to reduce the measurement error inherent in large-scale surveys and to report comparable performance scales because survey participants respond only to a subset of the assessment items. The scale of plausible values ranges from 0 to 500. A higher score signifies a higher measured proficiency. In practice, each estimation is repeated 10 times for each plausible value. The average coefficients and standard errors of the 10 estimations are reported.

(Mann *et al.* 1997).[14] Like most surveys aiming to measure socioemotional skills, the STEP Household Survey includes a battery of 24 self-reported items, designed by developmental and personality psychologists, that are mapped to the eight domains (skills) just listed. Each socioemotional skills score is the result of the aggregation of these predefined items (three on average for this survey). The response categories for each item range from 1, "almost never," to 4, "almost always." The aggregation of items onto domains is done through an inter-item average—that is, a weighted average of pre-assigned items based on all possible pairs.[15]

4. Empirical Strategy

Our objective is to investigate the distinct contribution of cognitive and socioemotional skills to labor market outcomes (i.e., labor earnings, labor force participation, and occupational choices) and tertiary education attainment. In other words, this study examines how those skills, however acquired, matter for these outcomes.

We use three different approaches in our analysis: a structural model, OLS and logit, and IV.

In view of our research question, we do not investigate the distinct effects of education and skill stocks on outcomes. Schooling, cognitive skills, and socioemotional skills are interrelated and influence each other. Distinguishing their respective effects requires a data structure different to the one available in STEP and can lead to severe estimation bias when considered simultaneously in a standard regression model (see Cawley, Heckman, and Vytlacil 2001; Heckman, Stixrud, and Urzúa 2006). Not controlling for schooling in linear estimations, such as a wage equation, can lead to overestimations of the net effects of measures of skills on wages but capture the whole effect of skills, independently of where they were formed (Heckman, Stixrud, and Urzúa 2006). We present both results controlling and not controlling for education, but we prioritize the approach excluding schooling for the interpretation.

4.a Reduced Form

The first empirical approach follows a standard Mincer-like specification to estimate the following relationship between a labor market or schooling outcome and a set of skills:

$$Y_i = \alpha + \beta_1^Y C_i + \beta_2^Y SE_i + \beta_3^Y X_i + \varepsilon_i \tag{1}$$

[14] See table 1 for brief definitions of the socio-emotional skills in the survey and items used for the construction of scores of these skills

[15] Empirical studies often perform exploratory factor analyses or the like to produce s*ocio-emotional* skills factors from the inventory of questions. However, the limited number of items available in the STEP survey prevented us from using such a method. The inter-item average approach has been empirically validated by lead psychologists who advised the World Bank STEP Core team and was performed with Stata's *alpha* command (World Bank 2014). Ultimately, it eases the interpretation of skill domain scores because items are preassigned to specific skills.

where Y_i is a labor market outcome (e.g., wage); C_i and SE_i represent, respectively, cognitive skills (such as reading proficiency) and socioemotional skills (such as conscientiousness) that affect the labor market outcome; and X_i is a set of factors (other than skills) that affect Y_i.

True skills C_i and SE_i are unobserved (latent). A common proxy for skills is years of schooling or school levels. However, educational attainment is a poor measure of actual scholastic ability because (1) many of the skills and personality traits that shape an individual's success are acquired outside the classroom, and (2) students acquire skills at each level of schooling very differently across schools and countries (Hanushek and Woessmann 2008). Nonetheless, the Colombia STEP Household Survey provides a set of measures, or test scores, T_i, that capture various dimensions of cognitive and socioemotional skills.

Assuming that T_i measures all skills that are captured in equation (1), we can rewrite the equation as

$$Y_i = \alpha + \beta_1^Y T_i^C + \beta_2^Y T_i^{SE} + \beta_3^Y X_i + \vartheta_i \tag{2}$$

Under the assumption that our sets of T_i perfectly measure all C_i and SE_i, we can estimate equation (2) using OLS or logit regressions without any ability bias, and β_1 and β_2 will give us the return to each skill captured by the vectors T_i^C and T_i^{SE}. However, a growing literature shows that measured skills, captured by the vector T_i, capture C_i and SE_i with error, so $\mathrm{Cov}(T_i, \vartheta_i) \neq 0$ could be possible (Borghans, Duckworth, Heckman, and ter Weel 2008; Hansen, Heckman, and Mullen 2004). In that case, measurement error and omitted variable bias produce biased estimates of β_1 and β_2.[16] The estimations featured in the next sections use the OLS and logit specifications as a benchmark for comparison with alternative methods.

4.b Structural Estimation[17]

An alternative to OLS and logit estimations to solve measurement error and omitted variable bias is to conduct a structural estimation of latent skills based on a measurement system of test scores (Keane and Wolpin 1997; Cameron and Heckman 2001; Heckman, Stixrud, and Urzúa 2006; Urzúa 2008; Sarzosa and Urzúa 2015, 2016). The outcomes of interest, Y, are a function of the latent skills and other factors influencing them, as depicted by the equation

$$Y = \alpha_A^Y \theta_A + \alpha_B^Y \theta_B + X_Y \beta^Y + e^Y \tag{3}$$

[16] Test scores are sensitive to the amount of schooling completed at the time of the test and family background (Hansen, Heckman, and Mullen 2004). Furthermore, the measures of ability are known to be very noisy. Thus using test scores as an independent variable in regression model analysis could lead to measurement error bias.
[17] For a more detailed explanation regarding identification and estimation see Sarzosa and Urzúa (2016).

where, θ_A and θ_B are the latent factors or dimensions of unobserved heterogeneity; β^Y, α_A, and α_B^Y are coefficients to estimate; X_Y is observable controls (e.g., gender, age); and e^Y is a vector of independently distributed error terms orthogonal to X_Y, θ_A, and θ_B.

The need for a structural estimation relies in the assumption that θ_A and θ_B are unobservable—that is, the measures or scores available in the data are only proxies of the true latent variables that we want to use for the estimation (Bartholomew, Knott, and Moustaki 2011). They are treated as realizations of the score-production function

$$T = X_T\beta^T + \alpha_A^T\theta_A + \alpha_B^T\theta_B + e^T \tag{4}$$

where T is an $L \times 1$ vector of scores (e.g., measures of reading proficiency, emotional stability, or grit); 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 takes e_Y and e_T as mutually independent conditional on (θ_A, θ_B, X).

Carneiro *et al.* (2003) and Sarzosa and Urzúa (2016) show that the system of production functions of test scores (4) can be used to non-parametrically identify the distributions of the latent abilities $f_{\theta_A}(\cdot)$ and $f_{\theta_B}(\cdot)$, their loading matrices (α_A and α_B), and the diagonal matrix of their variance, Σ_θ.[18] The loading matrices of latent factors, α_A and α_B, can be identified up to one normalization—that is, one loading per factor is set to equal to 1 and the rest of them will be interpreted relative to the one chosen as nummeraire. Kotlaski (1967) and Carneiro *et al.* (2003) show that two assumptions are needed for identification: (i) that latent skills factor θ_s for $s = \{A, B\}$ are orthogonal to each other, and (ii) that the system includes at least three test scores per skill. Estimating two factors of latent skills requires a minimum of six test scores ($L = 6$).

In practice, the test scores measurement system allows us to identify the distributions $f_{\theta_A}(\cdot)$ and $f_{\theta_B}(\cdot)$ associated to the unobserved heterogeneity in order to be able to integrate it away in a maximum likelihood procedure.[19] The likelihood function is then

$$\begin{aligned}\mathcal{L} = \prod_{i=1}^{N} \int\int f_{e^Y}(X_Y, Y, \varrho_1, \varrho_2) \times f_{e^{T_1}}(X_{T_1}, T_1, \varrho_1, \varrho_2) \cdots \\ \times f_{e^{T_6}}(X_{T_6}, T_6, \varrho_1, \varrho_2) dF_{\theta_1}(\varrho_1) dF_{\theta_2}(\varrho_2).\end{aligned} \tag{5}$$

---

[18] 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 re-create a wide range of distributions (Frühwirth-Schnatter 2006).

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

From which we 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 $f_{\theta_B}(\cdot)$.

4.c Robustness Checks: Exploring Instrumental Variable Estimations

In both conventional and advanced methods, the causal effects of our measures of skills on labor market outcomes can hardly be claimed because of the simultaneity of the observation of both the outcomes of interest and the measures of skills. Because we use cross-sectional data, reverse causality could be at play if a labor market outcome of interest would also influence its expected determinants such as measured skills (Carneiro and Heckman 2004). For example, although personality is often viewed as fairly stable over the life course, evidence from the United States and Germany suggests that participation in the labor market affects personality traits such as emotional stability, agreeableness, conscientiousness, and openness to experience (Gottschalk 2005; Boyce *et al.* 2015).

Instrumental variable methods provide consistent estimates even in the presence of measurement error and simultaneity (Angrist and Krueger 2001). This two-stage procedure relies on the use of an instrumental variable, Z_i, that is correlated with T_i^C and T_i^{SE} but not with the error term ϑ_i. For example, the proximity to school might be correlated with skills acquired through schooling but not with wages. The first stage estimates

$$T_i = \pi_0 + \pi_1^T Z_i + \pi_2^T X_i + u_i \tag{6}$$

and relates to the reduced-form equation as

$$Y_i = \pi_0 + \pi_1^Y Z_i + \pi_2^Y X_i + \varphi_i. \tag{7}$$

A trouble that arises when implementing an IV approach is the difficulty in selecting a meaningful instrument, highly correlated with T_i, and the need to select an IV for each T_i (many of which are produced through similar processes).[20] Invalid or weak instruments can lead to severe result bias (Murray 2006). For that reason, we use IV results as a robustness check—to inquire whether the relationships between labor outcomes and skill measures are generally consistent over estimation methods holding distinct sets of assumptions.

5. Descriptive Statistics

[20] For a discussion of these challenges, see Heckman, Stixrud, and Urzúa (2006), Heckman and Urzúa (2010), and the sources cited therein.

This section presents some descriptive statistics on the distribution of skills across gender, age groups, and educational levels of the working-age population. These distributions are not conditional on other observable and unobservable characteristics of individuals.

Reading proficiency levels are significantly different across educational level and generations but not gender. The distribution of scores is highly correlated (as expected) with educational level, though not perfectly (figure 1). In particular, the difference in mean levels between those with a secondary and tertiary education is not as pronounced as that between those with a primary and secondary education. On top of that, there are significant overlaps across education levels, suggesting that the completion of a schooling level does not necessarily guarantee a certain level of cognitive skills. The heterogeneity of reading proficiency within educational levels justifies the focus on individuals' skills, rather than educational achievement. While gender differences in reading proficiency are negligible, the distribution of scores among the young (15–24) is higher than among adults (25–49), a signal that suggests improvement over generations in the ability to understand and analyze written texts (also possibly correlated with educational levels) or that this ability tends to depreciate with aging.

In terms of socioemotional skills (figures 2, 3, and 4), differences across gender, age, and educational level are less noticeable. Across gender, and among all possible dimensions covered by the survey, men and women show slight differences in distribution of the conscientiousness, emotional stability, grit, decision-making, and hostile attribution bias scores. The most noticeable is the fact that males tend to score higher on the emotional stability scale than women, that is, reporting to be more able to manage emotions and stressful situations. By age, the only differences are registered for agreeableness, emotional stability, grit, and hostile attribution bias, with the young scoring lower than adults across these dimensions. Finally, there are significant differences in socioemotional skill scores by educational level; in all cases except for hostile attribution bias (which is the opposite), less educated workers score lower on the scale.

Correlations among measures of cognitive and socioemotional skills are often significant but rather modest. As shown in table 2, the correlation between reading proficiency and socioemotional dimensions differs substantially, with openness to experience, decision making, and hostile attribution bias among the ones with a higher correlation but never higher than 0.25. Some socioemotional dimensions are also relatively higher correlated among themselves—for example, 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 with grit (0.21).

6. Results

6.a OLS Estimates

Our first set of results explores OLS estimates of the relationship between disaggregated measures of cognitive and socioemotional skills and labor market outcomes. The first set of outcomes is for log hourly labor earnings (wage for salaried workers and net profits for self-employed), and these results are presented in table 3. The sample includes individuals between 15 and 64 years of age (both men and women).

Our main finding 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 labor earnings: an increase in one standard deviation in reading proficiency is correlated with a 15-percent increase in hourly labor earning from one's main job. As for socioemotional skills, only openness to experience seems to be significantly related to labor earnings: an increase in one standard deviation in openness to experience is correlated with an 8-percent increase in hourly labor earning from one's main job. These results remain when all skill dimensions are included in the same regression. It is important to note that the estimates in table 3 do not control for educational level, and so the coefficients capture the full association between different skills dimensions and labor earnings, irrespective of whether these skills were formed at school or at work.

Skills also have distinctive relationships with labor participation outcomes and occupational choices, including the likelihood of being a formal worker, of being a high-skilled worker, of being employed, of being active or studying, or of having pursued a tertiary education degree. Reading proficiency is again positively related with the probability of being a formal or high-skilled worker, but socioemotional skills seem to play no role in these outcomes—except for more hostile individuals having a lower probability of holding a formal job (table 4). However, some of these socioemotional characteristics seem to be relevant for labor or educational choice paths. For example, conscientiousness and decision-making are positively related with being employed and being active or in school (table 5). A higher scale in openness to experience, emotional stability, decision-making, and hostile attribution bias seem to matter for pursuing a tertiary education. For comparison purposes, tables 4 and 5 also include regressions that control for the educational level of the individual, the only difference being in the role of reading proficiency, which becomes nonsignificant (except for being a high-skilled worker), suggesting that educational level is the signal to which the job market responds and that it serves as a guarantee of the reading proficiency the employer is buying when hiring an individual with a given educational level.

Tables 6 and 7 present results for different gender, age, and education subgroups.[21] The main findings are 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—that is, those with at least a complete upper secondary education (nine years of schooling). By contrast, it is only related with labor force or school participation among females, young people (less than 35 years old), and less educated workers (maximum incomplete

[21] Only two outcomes are showcased: hourly earnings and being active or in school. Results for other labor market outcomes for subgroups are available upon request.

upper secondary education). As for socioemotional skills, the role of openness to experience in wages seem relevant only among males, older, and more educated workers. The role of socioemotional skills in explaining labor and schooling decisions also has different effects across subgroups. Strikingly, among males socioemotional skills do not play any role in occupational decisions.

6.b Structural Estimation

As stated earlier, we can also estimate the effect of latent skills on labor market outcomes using structural estimation methods as developed in Keane and Wolpin (1997); Carneiro *et al.* (2003); Heckman, Stixrud, and Urzúa (2006); and Sarzosa and Urzúa (2016). As opposed to the OLS and logit estimates presented earlier, this method can mitigate measurement error concerns because it acknowledges the fact that skills are latent rather than observable, abstracting from single (and potentially poor) measures of skills.

To apply this method, we construct an adjunct measurement system that comprises scores in two dimensions: cognitive skills and socioemotional skills. Identification in this set-up requires at least three test scores per dimension explored—that is, we had to construct three scores that provided information about socioemotional skills and three scores about cognitive skills.

To obtain the latent socioemotional skills factor, we aggregate the scale of extraversion with the measure of openness to experience into one score, the measure of emotional stability with the measure of hostile attribution bias into a second score, and the measures of conscientiousness, grit, and decision making into a third score. This aggregation secures the smoothness needed in the measurement system given that all of these measures come from categorical answers. The pairing of the measures is based on the correlations among them (table 2).

To obtain the measurement system needed to identify the factor of latent cognitive skills, we used the following scores: (1) a measure of language captured by a weighted average of reading components measured by the reading proficiency test—average print vocabulary, sentence processing, and passage comprehension (ETS 2014, World Bank 2014); (2) a measure of use and length of the reading undertaken on and off the workplace; and (3) a plausible value of reading proficiency, from the direct assessment, randomly chosen among 10.

Using these scores and exogenous controls such as age, gender, mother's education, and city of residence, we estimated the system of equations described in equations (3) and (4) of section 5.b.[22] The purpose of these estimations is to retrieve the components of the unobserved heterogeneity free of the exogenous characteristics that can affect the scores we observe.[23] These estimations are presented in tables 8 and 9. For example, people with more educated mothers are more likely

[22] All the estimations presented in this section were implemented using the *heterofactor* command in Stata developed by Sarzosa and Urzúa (2016).
[23] See figure 5 of the estimated distribution.

to have a broader vocabulary and thus score higher on the average print vocabulary, even if latent cognitive skills are unchanged.

More important than these coefficients are the estimated distributions of the unobserved heterogeneity obtained from these estimations. These distributions are used to structurally model the unobserved heterogeneity in the outcome equations. Figure 6 presents the variance decompositions of the scores. It shows that the latent factors explain large proportions of the variance of many of the scores. That variation is the one we identify as the latent skill or unobserved heterogeneity.

Having estimated the distributions that describe cognitive and socioemotional skills, we estimate their relationship with labor market outcomes. The results presented in table 10 indicate that the unobserved heterogeneity matters in almost every outcome we analyzed, yet in very different ways.[24] Socioemotional skills matter in choices such as participating in the labor market and attending college—that is, socioemotional skills prevent inactivity among those members of the population working and studying, although the probability of attending college is also highly correlated with cognitive skills. Once in the labor market, cognitive skills are the ones related with higher probabilities of formally working, being a high-skilled 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 hourly labor earnings.

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 of interest, at every level of the unobserved heterogeneity. Hence, we simulate the expected outcome as a function of this unobserved heterogeneity. Given that we estimate two dimensions of such heterogeneity, we present the simulations using three-dimensional graphs that plot

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

In that sense, we randomly draw θ_A and θ_B from the distributions, $f_{\theta_A}(\cdot)$ and $f_{\theta_B}(\cdot)$, estimated in the first-step estimations (described in tables 8 and 9) and construct $E[Y|\theta_A, \theta_B]$. This way, from the simulated graphs, we clearly see how the unobserved skills relate with the outcome variable.[25]

The probability of being active either in the labor force or as a student is higher as socioemotional and cognitive skills increase (Figure 7). Although the low-skilled population (both in the cognitive and socioemotional dimension) has a 78 percent probability of being active, the high-skilled population has a 95 percent probability of being so. If focused on only one dimension, we see that, other things being equal, a person gains 9 percentage points in the probability of being

[24] Consistent with the previous preferred specification, these estimations do not control for education.

[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 $\zeta \sim \mathcal{N}(0,1)$.

active if taken from the first to the 10th decile in the distribution of socioemotional skills. In the same way, an increase of 5.8 percentage points in the probability of being active is associated with taking a person from the first to the 10th decile of the distribution of cognitive skills.

Figure 7 reveals that the relationship between going to college and skills is even stronger. Those with the lowest levels of cognitive skills have almost no chance of going or having gone to college (only 1.5 percent), whereas 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 a relationship is dramatically different. Changing a person's socioemotional skills, from those available in the first decile to those available in the tenth decile of the distribution, is associated with an increase 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 in the 10th decile of the cognitive skill distribution, leaving everything else constant.

The size of these relationships 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 about 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 cognitive skills. Figures 10, 11, and 12 attest to this. For example, figure 9 shows that, all else being held constant, the likelihood of having a formal job increases by 28 percentage points (i.e., more than doubles) when a person in the first decile is compared with one in the 10th decile of the cognitive skill distribution. In the same way, figure 10 shows that workers with higher cognitive skills are 26.7 percentage points more likely to have a high-skilled job than low-skilled workers, who have a 28 percent probability of doing so. We also find that workers who belong to the top decile of the cognitive skill can earn up to Col$3,000 (US$1.50) more per hour than those in the lowest decile of the cognitive skills distribution, which is roughly 50 percent more (figure 12).

6.c IV Estimates

We complement our results from the OLS and structural estimations with results obtained from instrumental variables (IV) estimations, in an attempt to provide estimates free of measurement error, omitted variable bias, or reverse causality.

Among the possible instruments that could be appropriate and available in the Colombia STEP Household Survey, we select the age at which a person started school and the economic situation of the household the person lived in at age 12 (a self-reported variable based on a scale of from 1 to 10) as suitable instruments for the reading proficiency score. Reading proficiency is a skill that is fostered in school at a young age and to be influenced by a child's economic and family environment (Cawley, Heckman and Vytlacil 2001; Carneiro and Heckman 2004,). For each socioemotional skill, we used the indicator of whether the individual lived with both parents at age 12, and again the economic situation of the household at age 12. As for cognitive skills, socioemotional ones are greatly shaped by family and social environment (Cunha *et al.* 2006; Cunha *et al.* 2012).

First-stage results show that selected instruments significantly correlate in most cases with the skills variables to be instrumented. The instrument of cognitive skills correlate strongly with the cognitive variable: the age at which one started school is significantly and negatively correlated to reading proficiency and a better household's economic situation at age 12 is significantly positively correlated to reading proficiency (Table 11). The instruments for socioemotional skills do not correlate equally with all eight socioemotional skills measures: a higher household's economic situation at age 12 also drives higher levels of grit but also less conscientiousness, hostility bias, and careful decision making. Having lived with one or two parents at age 12 show no significant correlation with all socioemotional skills but for emotional stability.

The instruments for socioemotional skills are likely to be weak, which is not the case of the cognitive skills IV. The cognitive skills IV are above critical values of Stock, Wright, and Yogo (2002), are valid instruments (no over- or under-identification as per the Sargan-Hansen test of over-identified restrictions and under-identification test), and are significant in the structural equation (Table 12). By contrast, socioemotional skills IV are below critical values of Stock, Wright, and Yogo (2002), are not overidentified (per Sargan-Hansen test) but do not pass the under-identification test (meaning IV are weakly correlated with their endogenous regressors); they are significant in the structural equation. We expect the weakness of socioemotional skills IV to results from the lower variance of the variables to instrument compared to the one of cognitive skills.

The second-stage IV estimations confirms some of the results found with other methods (OLS, structural estimation of latent skills). Cognitive skills are strongly correlated with labor earnings (more than twice more than OLS estimations) and with attending tertiary education (three times more than OLS estimations). Socioemotional skills (conscientiousness, emotional stability, grit, and hostile attribution bias) are also strongly correlated with attending tertiary education (Table 13).

However, some unexpected results arise. No set of skills is significant for the probability of being a formal worker, of being employed, or being active or in school. By contrast, reading proficiency and several socioemotional skills (conscientiousness, emotional stability, grit, and hostile attribution bias) seem to be significantly related with the probability of being a high-skilled worker, which was not observed in OLS estimations in the case of socioemotional skills. While grit was not significantly correlated with outcomes in OLS estimations one increase of a standard deviation of it increase the likelihood of being a high-skilled worker of 33 percentage points and of attending tertiary education of 54 percentage points. Moreover, conscientiousness is negatively correlated with outcomes in OLS estimations while it the opposite in IV estimations.

Although these results should be interpreted with caution due to the relative weakness of socioemotional skills IV it confirms that cognitive skills are consistently correlated with outcomes of good-quality job and tertiary education and that various specific socioemotional skills are correlated with various outcomes.

7. Conclusion

Using a unique data set that measures cognitive (reading proficiency) and socioemotional skills (personality traits and behaviors) for Colombia, we have 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 of latent skills).

Across all methods, one result came up quite consistently: cognitive skills (in particular, reading proficiency) are an important predictor of earnings and quality job. For example, 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 high-skilled worker. This result is consistent with previous findings such as Murnane, Willett, and Levy (1995); Murnane *et al.* (2000); Altonji and Pierret (2001); Cawley, Heckman, and Vytlacil (2001); and Hanushek and Woessmann (2008) for the United States.

By contrast, the role of socioemotional skills seems quite different. Across all the 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 United States such as Bowles, Gintis, and Osborne (2001a, 2001b) and Drago (2011). But they do seem to play an important role as a predictor of labor force participation and schooling decisions, as found in Carneiro, Crawford, and Goodman (2007) and Almlund *et al.* (2011).

Some results differ by subgroups (because of data limitations, it was only possible to perform this breakdown using OLS methods). For example, socioemotional skills are better predictors of labor force participation among women, younger people (under 35), and less educated workers (less than a complete secondary education). By contrast, cognitive skills seem more relevant for explaining wage levels among males, older people, 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 socioemotional 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. Given the influence of family environments on skill formation, there is also a great role to play for parenting and extra-curricular activities to foster cognitive and socioemotional skills. In any case, further research is needed on the optimal combination of packages for different demographic and socioeconomic population groups, rarely systematically incorporated into the education system, particularly in developing countries.

References

Almlund, M., A. L. Duckworth, J. J. Heckman, and T. Kautz. 2011. “Personality Psychology and Economics.” In *Handbook of the Economics of Education* Vol. 4, ed. E. A. Hanushek. Amsterdam: North Holland.

Altonji, J. G., and C. R. Pierret. 2001. “Employer Learning and Statistical Discrimination.” *Quarterly Journal of Economics* 116 (1): 313–50.

Angrist, J. D., and A. B. Krueger. 2001. "Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments." *Journal of Economic Perspectives* 15 (4): 69–85.

Barrick, M. R., and M. K. Mount. 1991. “The Big Five Personality Dimensions and Job Performance: A Meta-Analysis.” *Personnel Psychology* 44 (1): 1–26.

Bartholomew D., M. Knott, and I. Moustaki. 2011. *Latent Variable Models and Factor Analysis: A Unified Approach*, 3rd ed., Wiley Series in Probability and Statistics. West Sussex, UK: Wiley.

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

Borghans, L., A. L. Duckworth, J. J. Heckman, and B. ter Weel. 2008. “The Economics and Psychology of Personality Traits.” *Journal of Human Resources* 34 (4): 972–1059.

Borghans, L., B. H. H. Golsteyn, J. J. Heckman, and J. E. Humphries. 2011. “Identification Problems in Personality Psychology.” *Personality and Individual Differences* 51 (3): 315–20.

Borghans, L., B. ter Weel, and B. A. Weinberg J. 2008. “Interpersonal Styles and Labor Market Outcomes.” *Human Resources* 43 (4): 815–58.

Borghans, L., B. ter Weel and B. A. Weinberg. 2013. “People Skills and the Labor Market Outcomes of Underrepresented Groups.” *CBP Discussion Paper* 253. Den Haag: CPB Netherlands Bureau for Economic Policy Analysis.

Bowles, S., H. Gintis, and M. Osborne. 2001a. "The Determinants of Earnings: A Behavioral Approach." *Journal of Economic Literature* 39 (4): 1137–76.

______. 2001b. “Incentive-Enhancing Preferences: Personality, Behavior, and Earnings.” *American Economic Review* 91 (2): 155–58.

Boyce, C. J., A. M. Wood, M. Daly, and C. Sedikides. 2015. “Personality Change Following Unemployment.” *Journal of Applied Psychology*. Advance online publication.

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 162.

Caliendo, M., D. A. Cobb-Clark, and A. Uhlendorff. 2010. "Locus of Control and Job Search Strategies." IZA Discussion Paper 4750, Institute for the Study of Labor (IZA), Bonn.

Cameron, S., and J. J. Heckman. 2001. "The Dynamics of Educational Attainment for Black, Hispanic, and White Males." *Journal of Political Economy* 109 (3): 455–99.

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* 44 (2): 361–422.

Carneiro, P., and J. J. Heckman. 2004. "Human Capital Policy." In *Inequality in America: What Role for Human Capital Policy?* ed. J. J. Heckman, A. B. Krueger, and B. M. Friedman. Cambridge, MA: MIT Press.

Cattell, R. B. 1987. *Intelligence: Its Structure, Growth, and Action.* New York: Elsevier Science.

Cawley, J., J. J. Heckman, and E. Vytlacil. 2001. "Three Observations on Wages and Measured Cognitive Ability." *Labour Economics* 8: 419–42.

Cobb-Clark, D. A., and M. Tan. 2011. "Noncognitive Skills, Occupational Attainment, and Relative Wages." *Labour Economics* 18 (1): 1–13.

Cunha, F., J. J. Heckman, L. Lochner, and D. Masterov. 2006. "Interpreting the Evidence on Life Cycle Skill Formation." In *Handbook of the Economics of Education*, 1, ed. E. A. Hanushek and F. Welch, 697–812. Amsterdam: North Holland.

Cunha, F., J. J. Heckman, and S. M. Schennach. 2012. "Estimating the Technology of Cognitive and Noncognitive Skill Formation." *Econometrica* 78 (4): 883–931.

Cunningham, W., and P. Villaseñor. 2016. "Employer Voices, Employer Demands, and Implications for Public Skills Development Policy Connecting the Labor and Education Sectors." *World Bank Research Observer* 31 (1): 102–34.

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." World Bank, Washington, DC.

DiPrete, T. A., and J. L. Jennings. 2012. "Social and Behavioral Skills and the Gender Gap in Early Educational Achievement." *Social Science Research* 41 (1): 1–15.

Dodge, K. A. 2003. "Do Social Information Processing Patterns Mediate Aggressive Behavior? In *Causes of Conduct Disorder and Juvenile Delinquency*, "ed. B. B. Lahey, T. E. Moffitt, and A. Caspi. New York: Guilford Press.

Drago, F. 2011. "Self-Esteem and Earnings." *Journal of Economic Psychology* 32 (3): 480–88.

Duckworth, A. L., and M. E. P. Seligman. 2005. "Self-Discipline Outdoes IQ in Predicting Academic Performance of Adolescents." *Psychological Science* 16 (2): 939–44.

Duckworth, A., C. Peterson, M. Matthews, and D. Kelly. 2007. "Grit: Perseverance and Passion for Long Term Goals." *Journal of Personality and Social Psychology* 92 (6): 1087–101.

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* 43 (6): 1428–46.

ETS (Educational Testing Services). 2014. *A Guide to Understanding the Literacy Assessment of the STEP Skills Measurement Survey.* Princeton, NJ: IEA-ETS Research Institute.

Finnie, R., and R. Meng. 2001. "Minorities, Cognitive Skills, and Incomes of Canadians." *Canadian Public Policy* 28 (2): 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* 43 (4): 886–920.

Frühwirth-Schnatter, S. 2006. "Finite Mixture and Markov Switching Models." *Psychometrika* 74 (3): 559–560.

Gallo, W. T., J. Endrass, E. H. Bradley, D. Hell, and S. V. Kasl. 2003. "The Influence of Internal Control on the Employment Status of German Workers." *Journal of Applied Social Science Studies* 123 (1): 71–82.

Glewwe, P. 2002. "Schools and Skills in Developing Countries: Education Policies and Socioeconomic Outcomes." *Journal of Economic Literature* 40(2): 436–82.

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

Gottfredson, L. S. 1997. "Why g Matters—The Complexity of Everyday Life." *Intelligence* 24 (1): 79–132.

Gottschalk, P. 2005. "Can Work Alter Welfare Recipients' Beliefs?" *Journal of Policy Analysis and Management* 24 (3): 485–98.

Guerra, N., K. Modecki, and W. Cunningham. 2014. "Social-emotional Skills Development across the Life Span: PRACTICE." Policy Research Working Paper 7123, World Bank, Washington, DC.

Hansen, K. T., J. J. Heckman, and K. J. Mullen. 2004. "The Effect of Schooling and Ability on Achievement Test Scores." *Journal of Econometrics* 121 (1–2): 39–98.

Hanushek, E. A. 2015. "Why Standard Measures of Human Capital are Misleading." *Korea Development Institute Journal of Economic Policy* 37 (2): 22–39.

Hanushek, E. A., G. Schwerdt, S. Wiederhold, and L. Woessmann. 2015. "Returns to Skills around the World: Evidence from PIAAC." European Economic Review 73 (2015): 103–130.

Hanushek, E. A., and L. Woessmann. 2008. "The Role of Cognitive Skills in Economic Development." *Journal of Economic Literature* 46 (3): 607–68.

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 and Management Strategy* 19 (4): 947–89.

Heckman, J. J., and T. Kautz. 2012. "Hard Evidence on Soft Skills." *Labour Economics* 19 (4): 451–64.

Heckman, J. J., and Y. Rubinstein. 2001. "The Importance of Noncognitive Skills: Lessons from the GED Testing Program." *American Economic Review* 91 (2): 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* 24 (3): 411–82.

Heckman, J. J., and S. Urzúa. 2010. "Comparing IV with Structural Models: What Simple IV Can and Cannot Identify." *Journal of Econometrics* 156 (1): 27–37.

Heineck, G., and S. Anger. 2010. "The Returns to Cognitive Abilities and Personality Traits in Germany." *Labour Economics* 17 (3): 535–46.

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

John, O. P., and S. Srivastava. 1999. "The Big Five Trait Taxonomy: History, Measurement and Theoretical Perspectives." In *Handbook of Personality: Theory and Research*, ed. L. A. Pervin and O. P. John. New York: Guilford Press.

Judd, K. L. 1998. "Numerical Methods in Economics." *Journal of Economic Dynamics and Control* 25 (8): 1263–71.

Keane, M. P., and K. I. Wolpin. 1997. "The Career Decisions of Young Men." *Journal of Political Economy* 105 (3): 473–522.

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* 47 (6): 795–99.

Kotlarski, I. (1967). "On Characterizing the Gamma and the Normal Distribution". *Pacific Journal of Mathematics*. 20(1):69–76.

Kuhn, P., and C. Weinberger. 2005. "Leadership Skills and Wages." *Journal of Labor Economics* 23 (3): 395–436.

Lazear, E. P. 2003. "Teacher Incentives." *Swedish Economic Policy Review* 10 (2): 179–214.

______. 2005. "Entrepreneurship." *Journal of Labor Economics* 23 (4): 649–80.

Levine, R., and Y. Rubinstein. 2013. "Smart and Illicit: Who Becomes an Entrepreneur and Does It Pay?" NBER Working Paper 19276, National Bureau of Economic Research, Cambridge, MA.

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* 3 (1): 101–28.

Linz, S. J., and A. Semykina. 2008. "Attitudes and Performance: An Analysis of Russian Workers." *Journal of Socio-Economics* 37 (2): 694–717.

Mann, L., P. Burnett, M. Radford, and S. Ford. 1997. "The Melbourne Decision Making Questionnaire: An Instrument for Measuring Patterns for Coping with Decisional Conflict." *Journal of Behavioral Decision Making* 10 (1): 1–19.

Manning, A., and J. Swaffeld. 2008. "The Gender Gap in Early-Career Wage Growth." *Economic Journal* 118 (530): 983–1024.

McIntosh, S., and A. Vignoles. 2001. "Measuring and Assessing the Impact of Basic Skills on Labour Market Outcomes." *Oxford Economic Papers* 53 (3): 435–81.

Mohanty, M. S. (2010). "Effects of Positive Attitude and Optimism on Employment: Evidence from the US Data." *Journal of Socio-Economics* 39 (2): 258-270.

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

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* 19 (4): 547–68.

Murnane, R. J., J. B. Willett, and F. Levy. 1995. "The Growing Importance of Cognitive Skills in Wage Determination." *Review of Economics and Statistics* 77 (2): 251–66.

Murray, Michael P. M. P. 2006. "Avoiding Invalid Instruments and Coping with Weak Instruments." *Journal of Economic Perspectives* 20 (4): 111–32.

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* 51 (2): 77–101.

Nyhus, E. K., and E. Pons. 2005. "The Effects of Personality on Earnings." *Journal of Economic Psychology* 26: 363–84.

OECD (Organisation for Economic Co-operation and Development). 2012. *Literacy, Numeracy and Problem Solving in Technology-Rich Environments: Framework for the OECD Survey of Adult Skills.* Paris: OECD.

______. 2013. *Technical Report of the Survey of Adult Skills (PIAAC).* Paris: OECD.

______. 2014. *PISA 2012 Results: What Students Know and Can Do—Student Performance in Mathematics, Reading and Science,* Vol. I, Rev. Ed. Paris: OECD Publishing.

______. 2015. *Skills for Social Progress: The Power of Social and Emotional Skills*, OECD Skills Studies. Paris: OECD Publishing.

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* 26 (6): 827–41.

Prada, M. 2013. "Beyond Smart and Sociable: Rethinking the Role of Abilities on Occupational Choices and Wages." University of Maryland, College Park.

Prada, M., and S. Urzúa. 2014. "One Size Does Not Fit All: The Role of Vocational Ability on College Attendance and Labor Market Outcomes." *NBER Working Papers* 20752. National Bureau of Economic Research, Cambridge, MA.

Psacharopoulos, G. and E. Velez (1992), "Schooling, Ability, and Earnings in Colombia, 1988", *Economic Development and Cultural Change*, Vol. 40, No. 3, pp. 629-643.

______. 2015. "Bullying Among Teenagers: The Role of Cognitive and Non-cognitive Skills." *NBER Working Paper* No 21631.

Sarzosa, M., and S. Urzúa. 2016. "Implementing factor models for unobserved heterogeneity in Stata: The *heterofactor* command." *Stata Journal* 16 (1): 197–228.

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* 86 (1): 162–73.

Segal, C. 2013. "Misbehavior, Education, and Labor Market Outcomes." *Journal of the European Economic Association* 11 (4): 743–79.

Stock, J. H., J. H. Wright, and M. Yogo. 2002. "A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments." *Journal of Business and Economic Statistics* 20 (4): 518–29.

UNESCO (United Nations Educational, Scientific and Cultural Organizati*on*. 2014. *Resultados comparados SERCE-TERCE (*Comparative results SERCE-TERCE*)*. Santiago: UNESCO.

Urzúa, S. 2008. "Racial Labor Market Gaps: The Role of Abilities and Schooling Choices." *Journal of Human Resources* 43 (4): 919–71.

Uysal, S. D., and W. Pohlmeier. 2011. "Unemployment Duration and Personality." *Journal of Economic Psychology* 32 (6), 980–992.

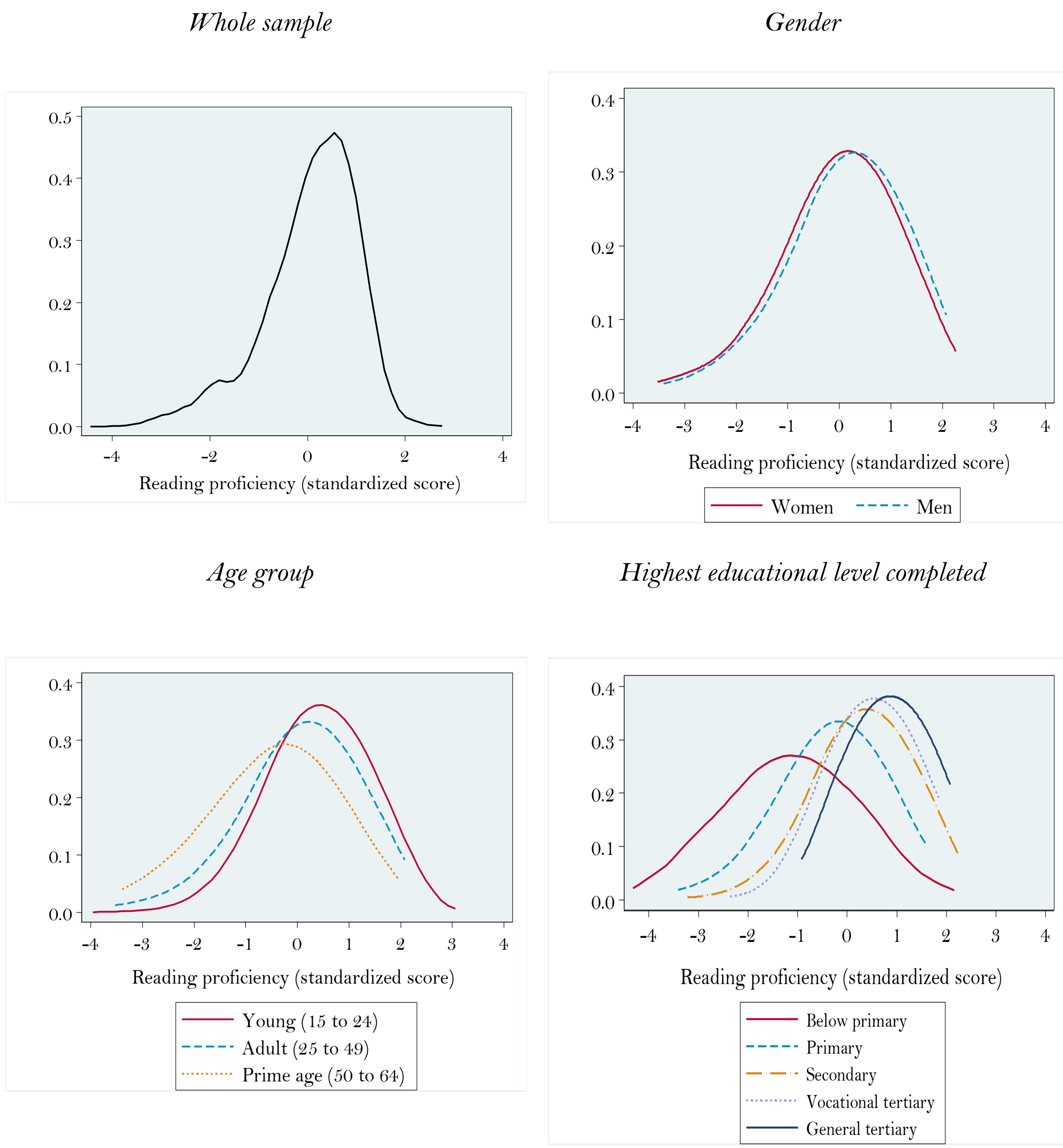
Von Davier, M., E. Gonzalez, and R. Mislevy. 2009. "What Are Plausible Values and Why Are They Useful?" In *IERI Monograph Series: Issues and Methodologies in Large Scale Assessments*, 2, ed. M. von Davier and D. Hastedt. Princeton, NJ: IEA-ETS Research Institute.

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

World Bank. 2014. "STEP Skills Measurement Surveys: Innovative Tools for Assessing Skills." Social Protection and Labor Discussion Paper 1421, World Bank, Washington, DC.

Figure 1 Distribution of Reading Proficiency across Groups of Interest, Colombia

Kernel densities of standardized reading proficiency scores

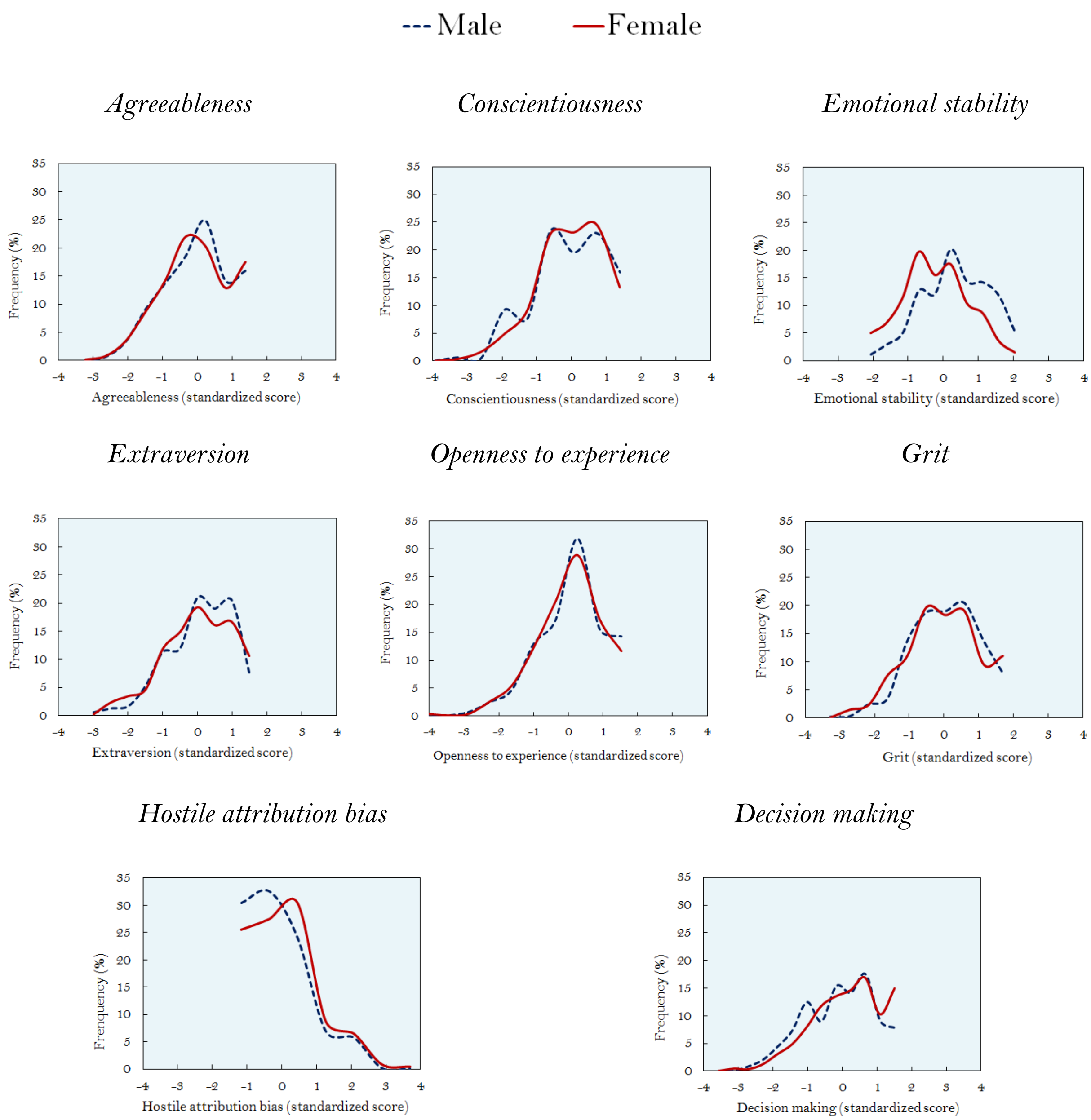


Source: Authors' calculations based on Colombia STEP Household Survey (2012).

Note: Differences in the distribution of reading proficiency scores are significant at the 95 percent level for age and educational levels (not across gender), based on two-sample Kolmogorov-Smirnov tests.

Figure 2 Distribution of Socioemotional Skills across Gender, Colombia

Share of individuals by socioemotional skills scores across gender

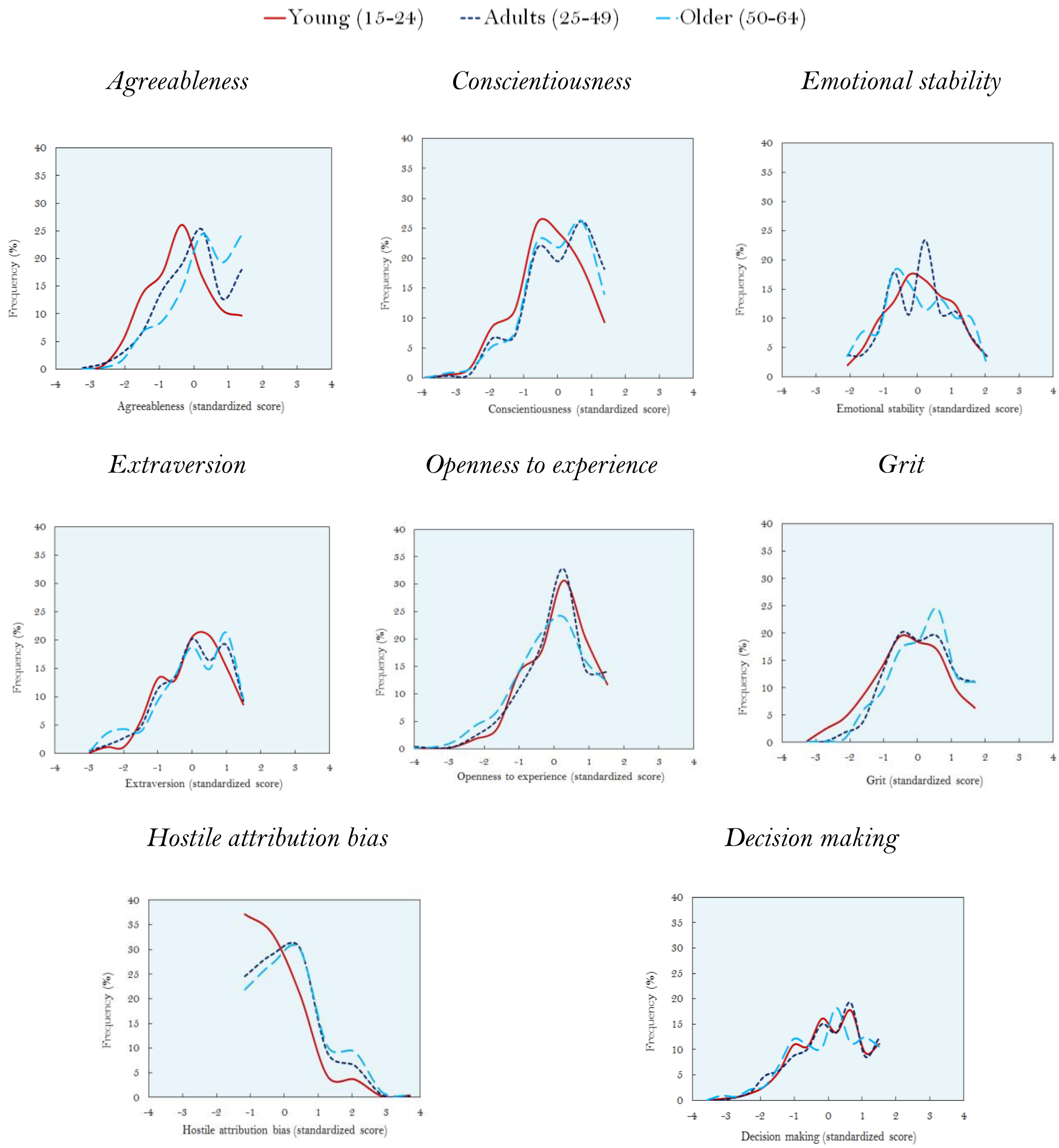


Source: Authors' elaboration based on Colombia STEP Household Survey (2012).

Note: The graphs represent scatter plots with smoothed lines based on tabulations of the standardized scores of socioemotional skills across gender. Differences in the distribution of conscientiousness, emotional stability, grit, decision making, and hostile attribution bias are significant at the 95 percent level based on Pearson's chi-square tests.

Figure 3 Distribution of Socioemotional Skills across Age Groups, Colombia

Share of individuals by socioemotional skills scores across age groups

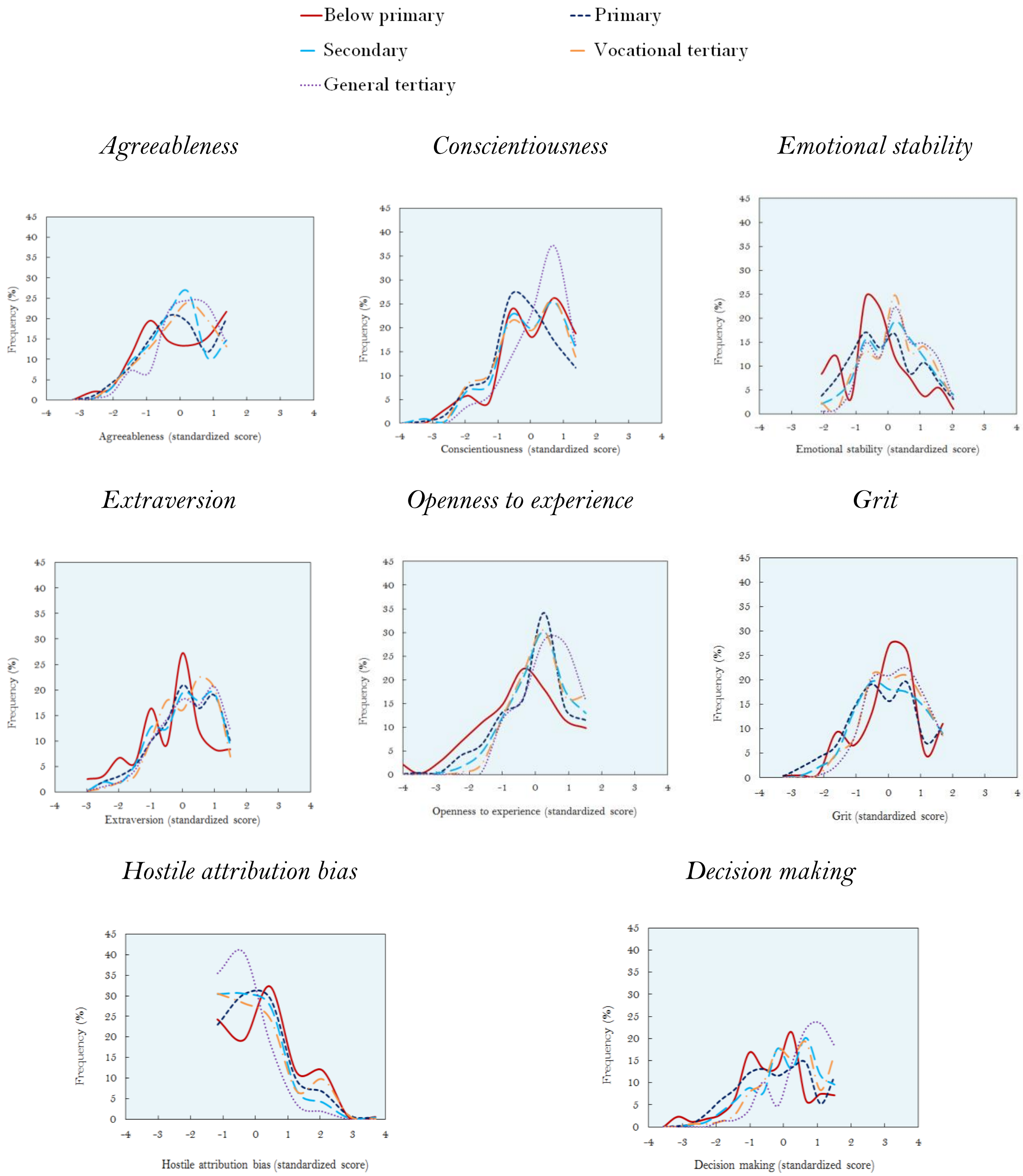


Source: Authors' elaboration based on Colombia STEP Household Survey (2012).

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

Figure 4 Distribution of Socioemotional Skills across Highest Educational Level Completed, Colombia

Share of individuals by socioemotional skills scores across educational levels

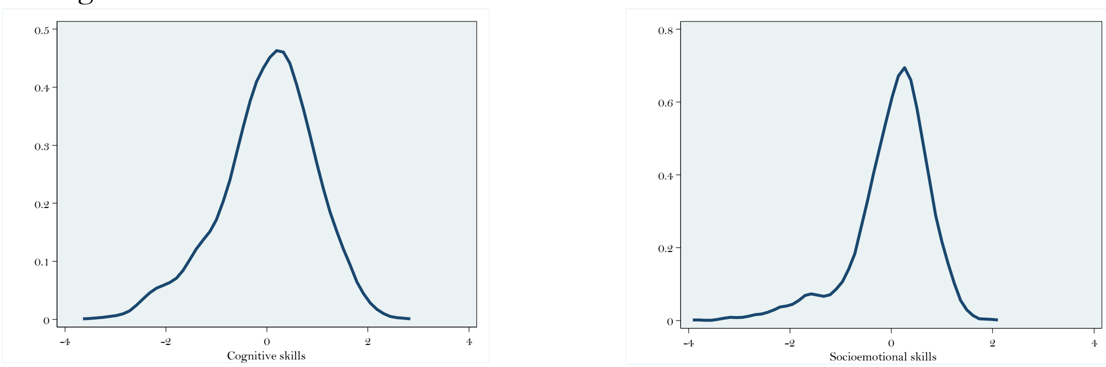


Source: Authors' elaboration based on Colombia STEP Household Survey (2012).

Note: The graphs represent scatter plots with smoothed lines based on tabulations of standardized scores of socioemotional skills across highest completed educational level. Differences in the distribution are all statistically significant at the 95 percent between at least at two levels based on Pearson's chi-square tests performed two levels by two.

Figure 5 Estimated Distribution of Cognitive and Socioemotional Skills Factors, Colombia

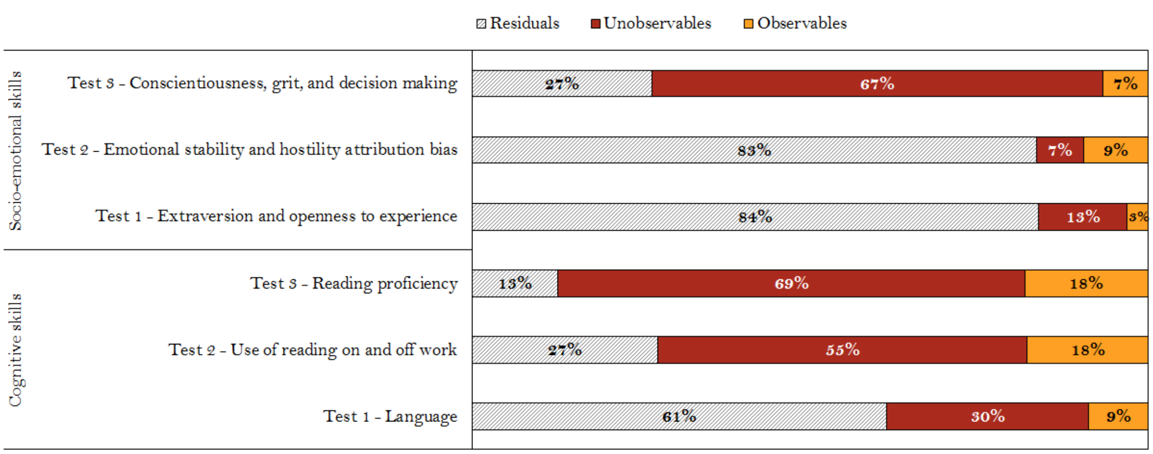
a. Cognitive skills b. Socioemotional skills



Source: Authors' calculations based on Colombia STEP Household Survey (2012).

Note: To obtain the measurement system identifying the factor of latent socioemotional skills, the scale of extraversion was aggregated with the measure of openness to experience (test 1), the measure of emotional stability with the measure of hostile attribution bias (test 2), and the measure of conscientiousness with the measures of grit and decision making (test 3)—see definitions in table 1. To obtain the measurement system identifying the factor of latent cognitive skills, we used the following tests: (1) a measure of languages captured by a weighted average of reading components measured by the reading proficiency test—average print vocabulary, sentence processing, and passage comprehension (World Bank 2014); (2) a measure of use and length of the reading done on and off the workplace; and (3) a plausible value of reading proficiency, from the direct assessment, randomly chosen among 10 (ETS 2014).

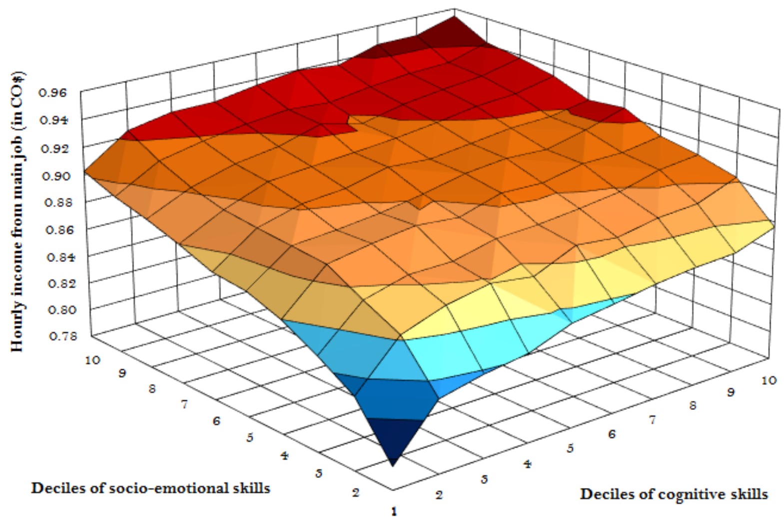
Figure 6 Variance Decomposition of the Tests Forming Socioemotional and Cognitive Skill Factors Used for Structural Estimation, Colombia



Source: Authors' calculations based on Colombia STEP Household Survey (2012).

Note: Factors of latent cognitive and socioemotional skills are obtained from a measurement system of three "test scores" for each. Measures of socioemotional skills were averaged into three tests to satisfy the necessary smoothness in the measurement system because all of these measures come from categorical answers; measures were paired based on the correlations among them (see table 2). To obtain the measurement system identifying the factor of latent socioemotional skills, the scale of extraversion was aggregated with the measure of openness to experience (test 1), the measure of emotional stability with the measure of hostile attribution bias (test 2), and the measure of conscientiousness with the measures of grit and decision making (test 3)—see definitions in table 1. To obtain the measurement system identifying the factor of latent cognitive skills, we used the following tests: (1) a measure of languages captured by a weighted average of reading components measured by the reading proficiency test—average print vocabulary, sentence processing, and passage comprehension (World Bank 2014); (2) a measure of use and length of the reading done on and off the workplace; and (3) a plausible value of reading proficiency, from the direct assessment, randomly chosen among 10 (ETS 2014).

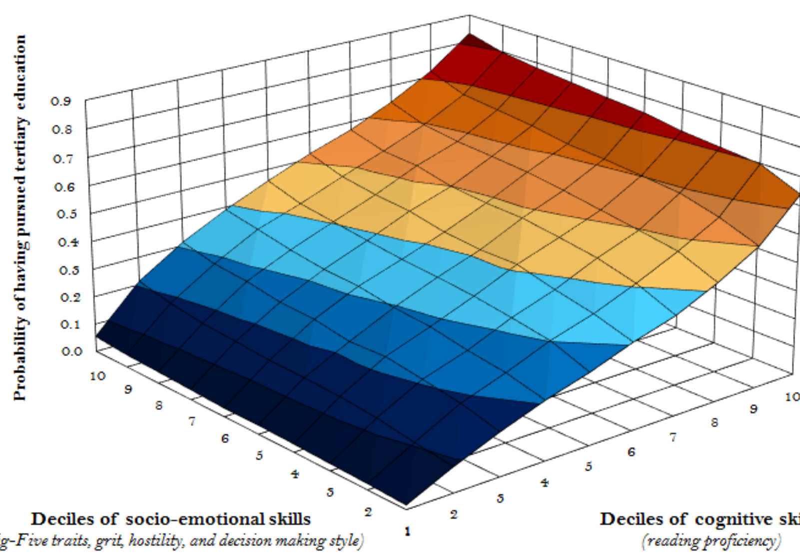
Figure 7 Probability of Being Active or in School by Skill Deciles, Colombia



Source: Authors' calculations based on Colombia STEP Household Survey (2012).

Note: Simulations are based on structural estimations of latent skills factors using Sarzosa and Urzúa (2016). Cognitive skills are captured by an estimated latent factor using measures of reading proficiency, use of reading on and off work, and language. Socioemotional skills are captured by an estimated latent factor using measures of personality traits (extraversion, openness to experience, emotional stability, conscientiousness, and grit) and behaviors (hostility bias and decision-making styles).

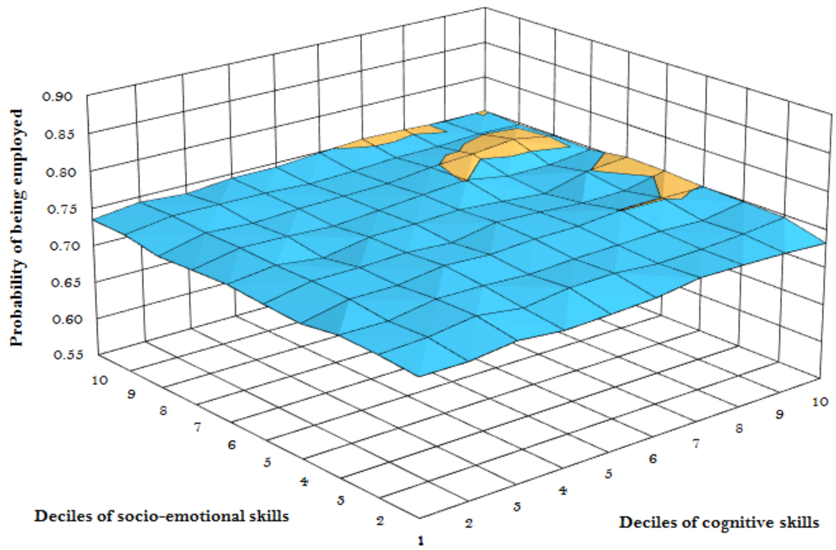
Figure 8 Probability of Having Pursued Tertiary Education by Skill Deciles for Adults Aged 25–64, Colombia



Source: Authors' calculations based on Colombia STEP Household Survey (2012).

Note: Simulations are based on structural estimations of latent skills factors using Sarzosa and Urzúa (2016). Cognitive skills are captured by an estimated latent factor using measures of reading proficiency, use of reading on and off work, and language. Socioemotional skills are captured by an estimated latent factor using measures of personality traits (extraversion, openness to experience, emotional stability, conscientiousness, and grit) and behaviors (hostility bias and decision-making styles).

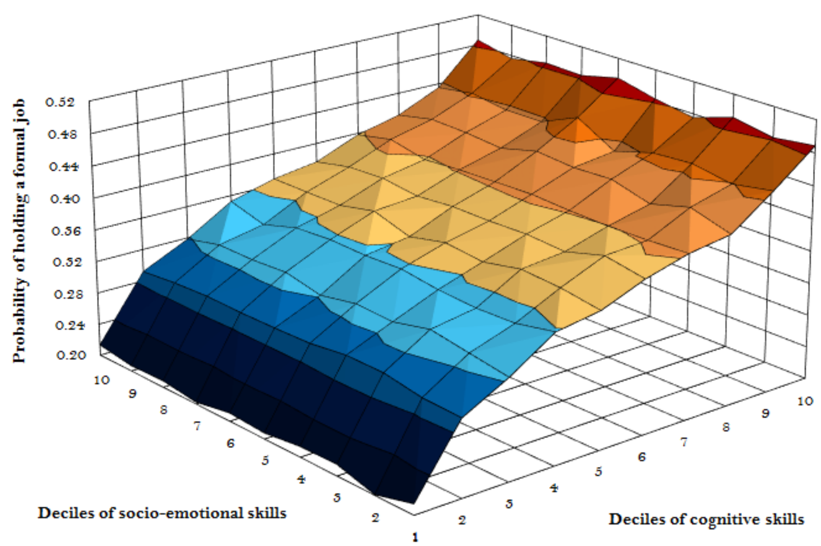
Figure 9 Probability of Being Employed by Skill Deciles for Adults Aged 19–64, Colombia



Source: Authors' calculations based on Colombia STEP Household Survey (2012).

Note: Simulations are based on structural estimations of latent skills factors using Sarzosa and Urzúa (2016). Cognitive skills are captured by an estimated latent factor using measures of reading proficiency, use of reading on and off work, and language. Socioemotional skills are captured by an estimated latent factor using measures of personality traits (extraversion, openness to experience, emotional stability, conscientiousness, and grit) and behaviors (hostility bias and decision making styles).

Figure 10 Probability of Holding a Formal Job by Skill Deciles, Colombia



Source: Authors' calculations based on Colombia STEP Household Survey (2012).

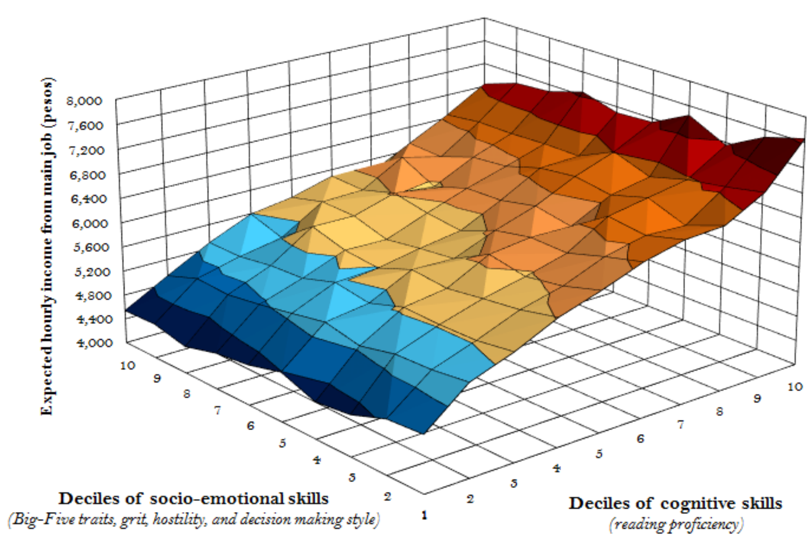
Note: Simulations are based on structural estimations of latent skills factors using Sarzosa and Urzúa (2016). Cognitive skills are captured by an estimated latent factor using measures of reading proficiency, use of reading on and off work, and language. Socioemotional skills are captured by an estimated latent factor using measures of personality traits (extraversion, openness to experience, emotional stability, conscientiousness, and grit) and behaviors (hostility bias and decision-making styles).

Figure 11 Probability of Holding a High-Skilled Occupation (versus Holding a Low- or Middle-Skilled Occupation) by Skill Deciles, Colombia

Source: Authors' calculations based on Colombia STEP Household Survey (2012).

Note: Simulations based on structural estimations of latent skills factors using Sarzosa and Urzúa (2016). Cognitive skills are captured by an estimated latent factor using measures of reading proficiency, use of reading on and off work, and language. Socioemotional skills are captured by an estimated latent factor using measures of personality traits (extraversion, openness to experience, emotional stability, conscientiousness, and grit) and behaviors (hostility bias and decision-making styles).

Figure 12 Hourly Income from Main Job by Skill Deciles: Colombia (pesos)



Source: Authors' calculations based on Colombia STEP Household Survey (2012).

Note: Simulations based on structural estimations of latent skills factors using Sarzosa and Urzúa (2016). Cognitive skills are captured by an estimated latent factor using measures of reading proficiency, use of reading on and off work, and language. Socioemotional skills are captured by an estimated latent factor using measures of personality traits (extraversion, openness to experience, emotional stability, conscientiousness, and grit) and behaviors (hostility bias and decision-making styles).

Table 1 Inventory of Socioemotional Skills in the Colombia STEP Household Survey

		Definition	Questionnaire item
Personality traits	Openness to experience	Appreciation for art, learning, unusual ideas, and variety of experience	Do you come up with ideas other people haven't thought of before?
			Are you very interested in learning new things?
			Do you enjoy beautiful things such as nature, art, and music?
	Conscientiousness	Tendency to be organized, responsible, and hardworking	When doing a task, are you very careful?
			Do you prefer relaxation more than hard work? R
			Do you work very well and quickly?
	Extraversion	Sociability, tendency to seek stimulation in the company of others, talkativeness	Are you talkative?
			Do you like to keep your opinions to yourself? Do you prefer to keep quiet when you have an opinion? R
			Are you outgoing and sociable—for example, do you make friends very easily?
	Agreeableness	Tendency to act in a cooperative, unselfish manner	Do you forgive other people easily?
			Are you very polite to other people?
			Are you generous to other people with your time or money?
	Emotional stability	Predictability and consistency in emotional reactions, with absence of rapid mood changes	Are you relaxed during stressful situations?
			Do you tend to worry? R
			Do you get nervous easily? R
	Grit	Perseverance with long-term goals	Do you finish whatever you begin?
			Do you work very hard? For example, do you keep working when others stop to take a break?
			Do you enjoy working on things that take a very long time (at least several months) to complete?
Behaviors and attitudes	Decision making	Manner in which individuals approach decision situations	Do you think about how the things you do will affect you in the future?
			Do you think carefully before you make an important decision?
			Do you ask for help when you don't understand something?
			Do you think about how the things you do will affect others?
	Hostile attribution bias	Tendency to perceive hostile intents in others	Do people take advantage of you?
			Are people mean/not nice to you?

Source: Authors' elaboration based on Almlund *et al.* (2011); John and Srivastava (1999); World Bank (2014).

Note: For each item, response categories range from 1 to 4: (1) almost never; (2) sometimes; (3) most of the time; (4) almost always. The score of each trait domain (e.g., extraversion) is the average of the individual scores on items of this trait. "R" refers to items that are reversely coded for the aggregation.

Table 2 Partial Correlations between Measures of Skills, Colombia

	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

Source: Authors' calculations based on Colombia STEP Household Survey (2012).

*$p < 0.001$.

Table 3 OLS Regressions of Log Hourly Labor Earnings on Cognitive Skills and Socioemotional Skills, Colombia

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

Source: Authors' calculations based on Colombia STEP Household Survey (2012).

Note: Standard errors are in parentheses. All regressions are estimated using OLS and include controls for being a woman (dummy), age, age-squared, mother's 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 1 percent of the log hourly labor earnings distribution are trimmed. Measures of reading proficiency and socioemotional skills are standardized. Regression coefficients and standard errors of reading proficiency are the average of the 10 estimations using plausible values.

*$p < 0.1$ **$p < 0.05$ ***$p < 0.01$

Table 4 OLS and Logit Regressions of Labor Earnings, Formality, and Occupational Status with Measures of Skills and Schooling, Colombia

Outcome	Log hourly labor earning		Being a formal worker		Being a high-skilled worker	
Method	OLS		Logit		Logit	
With/without schooling	Without	With	Without	With	Without	With
	(1)	(2)	(3)	(4)	(5)	(6)
Reading proficiency	0.161***	0.065	0.063***	0.014	0.141***	0.061***
	(0.05)	(0.06)	(0.02)	(0.02)	(0.02)	(0.02)
Extraversion	-0.009	0.000	0.001	0.005	-0.004	-0.000
	(0.04)	(0.04)	(0.02)	(0.02)	(0.01)	(0.01)
Conscientiousness	-0.034	-0.034	-0.003	-0.003	0.001	-0.002
	(0.04)	(0.04)	(0.02)	(0.02)	(0.01)	(0.01)
Openness to experience	0.082**	0.078**	-0.020	-0.022	0.020	0.017
	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)	(0.01)
Emotional stability	0.008	-0.015	0.027	0.016	0.006	-0.007
	(0.04)	(0.04)	(0.02)	(0.02)	(0.01)	(0.01)
Agreeableness	0.023	0.015	-0.011	-0.014	-0.007	-0.003
	(0.03)	(0.03)	(0.02)	(0.02)	(0.01)	(0.01)
Grit	-0.030	-0.043	-0.020	-0.025	0.013	0.007
	(0.04)	(0.04)	(0.02)	(0.02)	(0.01)	(0.01)
Hostile attribution bias	-0.003	0.023	-0.041**	-0.030*	-0.019	-0.003
	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)	(0.01)
Decision making	0.013	-0.007	0.012	0.002	0.019	-0.002
	(0.04)	(0.04)	(0.02)	(0.02)	(0.01)	(0.01)
Education: below primary		0.015		0.103		0.091
		(0.11)		(0.08)		(0.07)
Education: upper secondary		0.289***		0.198***		0.163***
		(0.10)		(0.05)		(0.04)
Education: vocational tertiary		0.371***		0.263***		0.278***
		(0.11)		(0.05)		(0.04)
Education: general tertiary		0.880***		0.348***		0.566***
		(0.15)		(0.06)		(0.05)
Observations	1,372	1,372	1,576	1,576	1,801	1,801
R-squared	0.11	0.16				

Source: Authors' calculations based on Colombia STEP Household Survey (2012).

Note: A worker is defined here as formal if he or she benefits from social security through his job. High-skilled workers hold occupations categorized as senior officials and managers, professionals, or technicians, as opposed to low- and middle-skilled workers such as clerks, service workers, machine operators, or laborers (jobless having held a job in the past year are also included). The classification is based on the International Labour Organization's 1988 International Standard Classification of Occupations (ISCO). Standard errors are in parentheses. Conditional correlations are computed from ordinary least squares (OLS) regressions for labor earnings and logit regressions for labor supply outcomes. The bottom and the top 1 percent of the log hourly labor earnings distribution are trimmed. OLS calculations control for being a woman (dummy), age, age-squared, mother's education (dummies; primary education is the reference category), and 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 one's education at the age of 12 (three levels). Average marginal effects are reported for logit regressions and reflect the changes in the probability of being observed in a labor or school participation situation with respect to the variables evaluated at the mean. Measures of reading proficiency and socioemotional skills are standardized. Regressions coefficients and standards errors of reading proficiency are the average of the 10 estimations using plausible values.

*$p < 0.1$ **$p < 0.05$ ***$p < 0.01$

Table 5 Logit Regressions of Employment, Activity, and Educational Trajectory with Measures of Skills and Schooling, Colombia

Outcome	Being employed		Being active or in school	Having pursued a tertiary education
Method	Logit		Logit	Logit
With/without schooling	Without	With	Without	Without
	(1)	(2)	(3)	(4)
Reading proficiency	0.003	-0.009	0.021*	0.199***
	(0.02)	(0.02)	(0.01)	(0.02)
Extraversion	-0.007	-0.007	0.009	-0.009
	(0.02)	(0.02)	(0.01)	(0.01)
Conscientiousness	0.044***	0.045***	0.023**	0.002
	(0.02)	(0.02)	(0.01)	(0.01)
Openness to experience	0.011	0.010	0.018*	0.045***
	(0.02)	(0.01)	(0.01)	(0.01)
Emotional stability	0.018	0.015	0.009	0.048***
	(0.02)	(0.02)	(0.01)	(0.01)
Agreeableness	-0.016	-0.016	-0.019*	0.001
	(0.02)	(0.02)	(0.01)	(0.01)
Grit	0.005	0.004	0.003	0.003
	(0.02)	(0.02)	(0.01)	(0.01)
Hostile attribution bias	-0.010	-0.008	-0.009	-0.049***
	(0.01)	(0.01)	(0.01)	(0.01)
Decision making	-0.031*	-0.033**	-0.003	0.055***
	(0.02)	(0.02)	(0.01)	(0.01)
Education: below primary		-0.063		
		(0.06)		
Education: upper secondary		0.003		
		(0.04)		
Education: vocational tertiary		0.053		
		(0.05)		
Education: general tertiary		0.066		
		(0.07)		
Observations	2,117	2,117	2,356	1,717

Source: Authors' calculations based on Colombia STEP Household Survey (2012).

Note: Standard errors are in parentheses. Conditional correlations are computed from logit regressions for labor supply and educational outcomes. Regressions control for being a woman (dummy), age, age-squared, mother's education (dummies; primary education is the reference category), cities of living and their metropolitan areas (dummies; Bogota-Barranquilla-Villavicencio is the reference category), and a self-reported categorical variable on parents' involvement in ones' education at the age of 12 (three levels). Average marginal effects are reported and reflect the changes in the probability of being observed in a labor or school participation situation with respect to the variables evaluated at the mean. Measures of reading proficiency and socioemotional skills are standardized. Regression coefficients and standard errors of reading proficiency are the average of the 10 estimations using plausible values.

*$p < 0.1$ **$p < 0.05$ ***$p < 0.01$

Table 6 OLS Regressions of Labor Earnings with Measures of Skills, across Subsamples, Colombia

Outcome	Log hourly labor earning					
Method	OLS					
Subsample	Men	Women	Younger (15–34)	Older (35–64)	Less educated (maximum, incomplete secondary)	More educated (minimum, complete secondary)
	(1)	(2)	(3)	(4)	(5)	(6)
Reading proficiency	0.160**	0.140**	0.134*	0.173***	0.019	0.180**
	(0.07)	(0.06)	(0.07)	(0.06)	(0.05)	(0.08)
Extraversion	-0.030	0.024	0.011	-0.011	0.007	0.008
	(0.05)	(0.05)	(0.05)	(0.04)	(0.04)	(0.05)
Conscientiousness	-0.037	-0.017	-0.102*	0.044	-0.040	-0.004
	(0.05)	(0.05)	(0.05)	(0.04)	(0.04)	(0.05)
Openness to experience	0.130***	0.046	0.030	0.152***	0.065	0.066*
	(0.04)	(0.05)	(0.04)	(0.05)	(0.04)	(0.04)
Emotional stability	-0.031	0.047	-0.013	0.036	0.011	-0.050
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
Agreeableness	0.028	0.013	-0.006	0.037	-0.003	0.050
	(0.04)	(0.05)	(0.04)	(0.04)	(0.04)	(0.04)
Grit	-0.033	-0.040	-0.040	-0.059	-0.015	-0.084
	(0.06)	(0.06)	(0.06)	(0.05)	(0.05)	(0.05)
Hostile attribution bias	-0.044	0.029	0.048	-0.053	0.036	-0.008
	(0.04)	(0.05)	(0.05)	(0.04)	(0.05)	(0.05)
Decision making	0.046	-0.030	0.033	0.019	-0.024	0.006
	(0.05)	(0.05)	(0.05)	(0.04)	(0.04)	(0.05)
Observations	686	686	678	694	438	934
R-squared	0.12	0.09	0.13	0.16	0.06	0.14

Source: Authors' calculations based on Colombia STEP Household Survey (2012).

Note: Standard errors are in parentheses. The bottom and top 1 percent of the log hourly labor earnings distribution are trimmed. Ordinary least squares (OLS) calculations control for being a woman (dummy), age, age-squared, mother's education (dummies; primary education is the reference category), and cities of living and their metropolitan areas (dummies; Bogota-Barranquilla-Villavicencio is the reference category). Measures of reading proficiency and socioemotional skills are standardized. Regression coefficients and standard errors of reading proficiency are the average of the 10 estimations using plausible values.

*$p < 0.1$ **$p < 0.05$ ***$p < 0.01$

Table 7 Logit Regressions of Being Active or in School with Skills, across Subsamples, Colombia

Outcome	Being active or in school (versus nonstudent inactive)					
Method	Logit					
Subsample	Men	Women	Younger (15–34)	Older (35–64)	Less educated (maximum, incomplete secondary)	More educated (minimum, complete secondary)
	(1)	(2)	(3)	(4)	(5)	(6)
Reading proficiency	-0.001	0.039**	0.030**	0.012	0.035**	0.015
	(0.01)	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)
Extraversion	-0.002	0.017	0.023***	-0.016	0.029*	0.003
	(0.01)	(0.02)	(0.01)	(0.02)	(0.02)	(0.01)
Conscientiousness	0.003	0.043***	0.020**	0.027*	0.046***	0.004
	(0.01)	(0.02)	(0.01)	(0.02)	(0.02)	(0.01)
Openness to experience	0.004	0.029*	0.022**	0.027*	0.010	0.016
	(0.01)	(0.02)	(0.01)	(0.02)	(0.02)	(0.01)
Emotional stability	-0.008	0.021	0.007	0.017	0.007	0.011
	(0.01)	(0.02)	(0.01)	(0.02)	(0.02)	(0.01)
Agreeableness	-0.011	-0.027*	-0.010	-0.029*	-0.045***	-0.002
	(0.01)	(0.02)	(0.01)	(0.02)	(0.02)	(0.01)
Grit	0.006	0.003	-0.003	0.012	0.004	0.006
	(0.01)	(0.02)	(0.01)	(0.02)	(0.02)	(0.01)
Hostile attribution bias	-0.009	-0.012	-0.006	-0.009	0.018	-0.018**
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)
Decision making	-0.013	-0.000	-0.005	0.002	-0.000	-0.007
	(0.01)	(0.02)	(0.01)	(0.02)	(0.02)	(0.01)
Observations	933	1,369	1,233	1,123	864	1,492

Source: Authors' calculations based on Colombia STEP Household Survey (2012).

Note: Standard errors in parentheses. Regressions control for being a woman (dummy), age, age-squared, mother's education (dummies; primary education is the reference category), cities of living and their metropolitan areas (dummies; Bogota-Barranquilla-Villavicencio is the reference category), and a self-reported categorical variable on parental involvement in one's education at the age of 12 (three levels). Average marginal effects are reported and reflect the changes in the probability of being observed in a labor or school participation situation with respect to the variables evaluated at the mean. Measures of reading proficiency and socioemotional skills are standardized. Regression coefficients and standard errors of reading proficiency are the average of the 10 estimations using plausible values.

*$p < 0.1$ **$p < 0.05$ ***$p < 0.01$

Table 8 Estimation of Socioemotional Skills Factor for Structural Estimation, Colombia

	Test 1	Test 2	Test 3
Measure of skills included in the test	Extroversion and openness to experience	Emotional stability and hostile attribution bias	Conscientiousness, grit, and decision making
Socioemotional skills	0.4412**	0.0447	1
	(0.174)	(0.047)	.
Age	0.0183*	-0.0200**	0.1009***
	(0.011)	(0.009)	(0.010)
Age-squared	-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)
Constant	9.2485***	1.7043***	7.6960***
	(0.197)	(0.169)	(0.189)

Source: Authors' calculations based on Colombia STEP Household Survey (2012).

Note: Factors of latent socioemotional skills are obtained from the measurement system of the three "test scores" presented in the table. Measures of socioemotional skills were averaged into three tests to satisfy the necessary smoothness in the measurement system because all of these measures come from categorical answers. The measures were paired based on the correlations among them (see table 2). The scale of extraversion was aggregated with the measure of openness to experience (test 1), the measure of emotional stability with the measure of hostile attribution bias (test 2), and the measures of conscientiousness with grit and decision making (test 3)—(see definitions in table 1). Standard errors are in parentheses. All the estimations include city dummies (coefficients not reported). "Mother ed." refers to mother's education. The omitted category of the mother's education variable is tertiary education and beyond. $N = 2{,}372$.

$^{*}p < 0.1$ $^{**}p < 0.05$ $^{***}p < 0.01$

Table 9 Estimation of Cognitive Skills Factor for Structural Estimation, Colombia

	Test 1	Test 2	Test 3
Measure of skills included in the test	Language	Use of reading on and off work	Reading proficiency
Cognitive skills	1.6001***	0.9055***	1
	(0.057)	(0.020)	.
Age	0.0626***	0.0100	0.0240***
	(0.019)	(0.007)	(0.007)
Age-squared	-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)
Constant	25.2210***	0.6072***	0.3168***
	(0.346)	(0.129)	(0.121)

Source: Authors' calculations based on Colombia STEP Household Survey (2012).

Note: Factors of latent cognitive skills are obtained from the measurement system of three "test scores" presented in the table. The following tests were used: (1) a measure of languages captured by a weighted average of reading components measured by the reading proficiency test—average print vocabulary, sentence processing, and passage comprehension (World Bank 2014); (2) a measure of use and length of the reading done on and off the workplace; and (3) a plausible value of reading proficiency, from the direct assessment, randomly chosen among 10. Standard errors are in parentheses. All the estimations include city dummies (coefficients not reported). "Mother ed." refers to mother's education. The omitted category of the mother's education variable is tertiary education and beyond. $N = 2{,}340$.

*$p < 0.1$ **$p < 0.05$ ***$p < 0.01$

Table 10 Structural Estimates of Associations between Labor Market Outcomes on Latent Skills Factors, Colombia

	Log hourly labor earning	Being formal worker	Being a high-skilled worker	Being employed	Being active or in school	Having pursued a tertiary education
	(1)	(2)	(3)	(4)	(5)	(6)
Cognitive skills	0.134***	0.276***	0.252***	0.023	0.112**	0.988***
	(0.032)	(0.052)	(0.04)	(0.042)	(0.047)	(0.076)
Socioemotional skills	-0.026	-0.004	0.046	0.013	0.143***	0.170***
	(0.028)	(0.044)	(0.035)	(0.04)	(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)
Age-squared	-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)	(0.18)	(0.145)	(0.167)	(0.232)	(0.267)
Mother ed. = primary	-0.507***	-0.14	-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)	(0.15)	(0.114)	(0.142)	(0.212)	(0.236)
Parental involvement—medium		0.07	-0.058	-0.03	0.004	0.1
		(0.093)	(0.075)	(0.082)	(0.093)	(0.096)
Parental involvement—strong		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

Source: Authors' calculations based on Colombia STEP Household Survey (2012).

Note: Standard errors are in parentheses. Estimated using Sarzosa and Urzúa (2016). The estimations presented in columns (2) t0 (6) assume a probit structure, hence the presented coefficients not marginal effects. All the estimations include city dummies (coefficients not reported). "Parental involvement" refers to a parent's regularity in checking a primary student's grades and exams (reference category is "no, never, or almost never"). "Medium" means "yes, sometimes," and "strong" means "yes, always, or almost always." A worker is defined here as formal if he or she benefits from social security through his or her job. High-skilled workers hold occupations categorized as senior officials and managers, professionals, or technicians, as opposed to low- and middle-skilled workers such as clerks, service workers, machine operators, or laborers (jobless having held a job in the past year are also included). The classification is based on the International Labour Organization's 1988 International Standard Classification of Occupations (ISCO).

*$p < 0.1$ **$p < 0.05$ ***$p < 0.01$

Table 11 First-Stage IV Regressions of Measures of Skills on Instruments, Colombia

VARIABLES	(1) Read.	(2) Extr.	(3) Cons.	(4) Open.	(5) Emot.	(6) Agre.	(7) Grit	(8) Host.	(9) Deci.
Age started school	-0.070**								
	(0.03)								
Household eco. situation age 12	0.053***	-0.015	-0.063**	-0.007	0.032	-0.005	0.083**	-0.055**	-0.050*
	(0.02)	(0.02)	(0.03)	(0.02)	(0.03)	(0.02)	(0.03)	(0.02)	(0.03)
Lived with one parent at age 12		-0.088	0.124	-0.114	-0.180	-0.078	0.067	-0.115	0.110
		(0.23)	(0.16)	(0.21)	(0.22)	(0.15)	(0.18)	(0.17)	(0.28)
Lived with both parents at age 12		-0.101	-0.029	0.028	-0.357*	0.013	0.049	-0.150	0.193
		(0.20)	(0.15)	(0.20)	(0.20)	(0.14)	(0.17)	(0.16)	(0.26)
Age	0.045***	-0.003	0.057***	-0.017	0.023	0.019	0.069**	0.022	0.024
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)	(0.03)
Age-squared	-0.001***	0.000	-0.001**	0.000	-0.000	-0.000	-0.001*	-0.000	-0.000
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Woman	-0.058	-0.060	0.059	0.065	-0.460***	-0.002	-0.070	0.083	0.323***
	(0.06)	(0.09)	(0.09)	(0.07)	(0.09)	(0.07)	(0.10)	(0.07)	(0.10)
Medellin	0.141*	-0.208**	-0.183	-0.011	-0.176*	0.098	0.558***	-0.330***	0.089
	(0.08)	(0.10)	(0.13)	(0.10)	(0.10)	(0.09)	(0.11)	(0.10)	(0.12)
Cali	0.173*	-0.138	0.236**	-0.188*	-0.263*	-0.025	-0.156	-0.246***	-0.231*
	(0.09)	(0.14)	(0.11)	(0.11)	(0.14)	(0.10)	(0.14)	(0.09)	(0.13)
Bucaramanga	0.288***	-0.035	-0.034	-0.204	-0.098	-0.050	-0.261*	-0.134	-0.307**
	(0.09)	(0.12)	(0.13)	(0.13)	(0.18)	(0.14)	(0.16)	(0.18)	(0.14)
Manizales	0.249*	0.167	0.195	0.338**	-0.010	0.228	0.108	-0.150	-0.163
	(0.14)	(0.14)	(0.16)	(0.16)	(0.19)	(0.17)	(0.14)	(0.13)	(0.17)
Cucuta	0.106	-0.192	-0.173	-0.132	-0.228*	-0.156	-0.122	0.188	-0.293**
	(0.09)	(0.13)	(0.13)	(0.13)	(0.13)	(0.14)	(0.17)	(0.14)	(0.14)
Ibague	0.282**	0.278**	0.244*	0.295**	0.126	0.260**	0.179	-0.092	0.220
	(0.11)	(0.13)	(0.13)	(0.15)	(0.14)	(0.12)	(0.17)	(0.12)	(0.14)
Mother edu.: primary	0.415***	0.006	0.008	0.115	0.279**	0.346***	0.021	-0.236*	0.327**
	(0.10)	(0.13)	(0.13)	(0.12)	(0.13)	(0.12)	(0.13)	(0.12)	(0.14)
Mother edu.: secondary	0.622***	0.058	0.013	0.006	0.430***	0.347***	-0.087	-0.213	0.300*
	(0.11)	(0.14)	(0.15)	(0.15)	(0.15)	(0.13)	(0.15)	(0.14)	(0.17)
Mother edu.: tertiary	1.081***	0.296	0.062	0.224	0.512***	0.575***	0.164	-0.169	0.820***
	(0.13)	(0.18)	(0.20)	(0.18)	(0.19)	(0.17)	(0.21)	(0.20)	(0.18)
Constant	-0.719*	0.212	-0.807*	0.330	-0.227	-0.970***	-1.867**	0.023	-0.750
	(0.38)	(0.49)	(0.45)	(0.46)	(0.53)	(0.37)	(0.77)	(0.42)	(0.64)
Observations	1,364	1,371	1,371	1,371	1,371	1,371	1,371	1,371	1,371
R-squared	0.24	0.03	0.06	0.03	0.10	0.07	0.13	0.07	0.06

Table 12 Tests of the validity of instrument variables

<table>
<tr><th></th><th>(1)</th><th>(2)</th><th>(3)</th><th>(4)</th><th>(5)</th><th>(6)</th><th>(7)</th><th>(8)</th><th>(9)</th></tr>
<tr><th>VARIABLES</th><th>Read.</th><th>Extr.</th><th>Cons.</th><th>Open.</th><th>Emot.</th><th>Agre.</th><th>Grit</th><th>Host.</th><th>Deci.</th></tr>
<tr><td>Instruments variables</td><td>1. Age at which a person started school (continuous, 5-12)
2. Self-reported economic situation of the household the person lived in at age 12 (continuous, 1-10)</td><td colspan="8">1. Economic situation of family at age 12 (continuous, 1-10)
2. If lived with 0, 1, or 2 parents at age 12 (dummies)</td></tr>
<tr><td colspan="10">Testing the strength of the instruments</td></tr>
<tr><td>F-statistic of excluded instruments in first stage</td><td>19.55</td><td>0.24</td><td>3.39</td><td>0.83</td><td>1.91</td><td>0.48</td><td>2.86</td><td>0.24</td><td>3.39</td></tr>
<tr><td>p-value</td><td>0.000</td><td>0.867</td><td>0.018</td><td>0.479</td><td>0.125</td><td>0.694</td><td>0.036</td><td>0.036</td><td>0.867</td></tr>
<tr><td colspan="10">Testing for underidentification and instrument redundancy</td></tr>
<tr><td>Overidentification test of all instruments (Hansen J statistic)</td><td>0.09</td><td>0.39</td><td>0.45</td><td>2.81</td><td>2.94</td><td>2.84</td><td>0.10</td><td>0.29</td><td>0.76</td></tr>
<tr><td>p-value</td><td>0.770</td><td>0.389</td><td>0.798</td><td>0.245</td><td>0.230</td><td>0.242</td><td>0.953</td><td>0.865</td><td>0.686</td></tr>
<tr><td>Underidentification test (Kleibergen-Paap rk LM statistic)</td><td>31.61</td><td>0.72</td><td>9.47</td><td>2.50</td><td>4.86</td><td>1.45</td><td>8.02</td><td>6.49</td><td>0.76</td></tr>
<tr><td>p-value</td><td>0.000</td><td>0.869</td><td>0.024</td><td>0.475</td><td>0.182</td><td>0.694</td><td>0.046</td><td>0.090</td><td>0.686</td></tr>
<tr><td colspan="10">Testing for the significance of the endogenous regressors in the structural equation</td></tr>
<tr><td>Anderson-Rubin Wald test</td><td>6.53</td><td colspan="8">3.22</td></tr>
<tr><td>p-value</td><td>0.038</td><td colspan="8">0.359</td></tr>
</table>

Table 13 IV Estimates (Second-Stage) of Associations between Labor Market Outcomes and Skills, Colombia

	Log hourly labor earning	Being formal worker	Being a high-skilled worker	Being employed	Being active or in school	Having pursued a tertiary education
	(1)	(2)	(3)	(4)	(5)	(6)
Reading proficiency	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

Source: Authors' calculations based on Colombia STEP Household Survey (2012).

Note: Standard errors are in parentheses. The bottom and the top 1 percent of the log hourly labor earnings distribution are trimmed. The instrumental variables (IVs) considered for reading proficiency are the age at which a person started school and the economic situation of the household at age 12; for each socioemotional 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 the Sargan-Hansen test. Other controls are gender, age, age-squared, mother's education (dummies; primary education is the reference category), and cities of living and their metropolitan areas (dummies; Bogota-Barranquilla-Villavicencio is the reference category). Measures of reading proficiency and socioemotional skills are standardized. Solely for the IV analysis, only the means of the 10 plausible values is used as a score of reading proficiency. For presentational purposes, the coefficients from different regressions for same outcome are presented in one column.

*$p < 0.1$ **$p < 0.05$ ***$p < 0.01$