

# Education, skills and a good job: a multidimensional econometric analysis

Jaya Krishnakumar and Ricardo Nogales

## 1 Introduction

There is a vast academic literature on the relation between labor market outcomes and education (see e.g. a literature review by [Goldberg & Smith \(2008\)](#)). A large number of empirical studies on this subject build upon the classic works of [Mincer \(1974\)](#) and [Becker \(1964\)](#), with success in the labor market being traditionally measured by earnings, which in turn are *directly* associated with educational achievements such as years of schooling (see e.g., [Heckman et al. \(2003\)](#)). In this study, we take a stand that a) the direct relation between labor market outcomes and wages should be qualified to include skills and abilities, and b) labor market success can hardly be described by a single measure. In what follows we elaborate on these two points, arguing for the need to go beyond traditional approaches to understand the relationship between education and a wide notion of work-related well-being.

The most dominant economic paradigm explaining labor market success, particularly in developing countries, can be represented by Mincer's equation (or variants of it), which relates earnings to years of schooling. However, years of schooling or more generally, time spent in formal educational programs, reflect *first-order* educational attainments that influence other characteristics such as abilities or skills ([Fasih, 2008](#)). Even if the time dedicated to formal education gives important signals to employers about a person's skills, it does not account for how these years have been effectively converted into skills by different individuals. It is these latter aspects of human capital that may transform into productivity and generate labor market returns.

The *technology of skill formation* pioneered and developed in [Cunha & Heckman \(2007, 2008\)](#); [Cunha et al. \(2010\)](#) incorporates the above ideas, explicitly accounting for skills as mediators in the relationship between educational investments and labor market

outcomes in adulthood. It has inspired a vast amount of applications using different outcome variables as the evaluative space of labor market outcomes (see e.g. [Lin et al. \(2018\)](#) for the US, [Lindqvist & Vestman \(2011\)](#) for Sweden and [Brunello & Schlotter \(2011\)](#) for a set of European countries). In this paper, we propose to use as the outcome of interest a *latent* work-related wellbeing variable that encompasses multiple aspects of a good job.

[Cunha & Heckman \(2007\)](#)'s framework is flexible enough to allow for a wide definition of lifestyle outcomes in adulthood such as earnings or employment status (see e.g. [Heckman et al. \(2011, 2018\)](#)). This is particularly useful for the study of *human development*, the definition of which is largely inspired from the Capability Approach (cf. [Sen \(1980, 1985, 1999\)](#)). Several international initiatives such as the UNDP's Human Development Index ([UNDP, 1990](#)), UN's Sustainable Development Goals ([UN, 2015](#)), OECD's Better Life Index<sup>1</sup>, European Commission's Going Beyond GDP<sup>2</sup> have managed to convince policymakers around the world to adopt a broad vision of *human* well-being covering many dimensions and including manifold aspects within each dimension. Based on the same theoretical underpinning, we consider well-being in the work dimension (work-related well-being), which is itself one of the many dimensions of human well-being ([Leßmann, 2012; Lugo, 2007](#)), as a *multifaceted* notion going beyond earnings or any single indicator. The concept of Decent Work launched by ILO ([Ghai, 2008](#)) also fits in with this multifaceted vision as it defines a good job as having many characteristics in addition to earnings such as job stability, social protection, low physical stress and so on. In a similar spirit, the OECD brings job quality to the forefront of the policy debate, while considering three essential dimensions of workers' well-being: earnings quality, labor market security and quality of the work environment ([OECD, 2014](#)). In spite of such initiatives and emphasis on a broad definition of a good job, studies relating education to such a multifaceted concept of a good job are still rare in the labor market literature. One can mention, for example, [Muñoz de Bustillo et al. \(2011\)](#), which posits that a better account of job quality should combine objective elements (e.g. earnings) with a subjective valuation of job attributes by the

workers themselves.

The empirical operationalization of the technology of skill formation in its original formulation is a challenging endeavor as it is heavily demanding in terms of information. Longitudinal datasets at the individual level, covering early childhood to adolescence/adulthood, with extensive information on the type of investments made by parents on children’s education, constitute an ideal database for its practical application. This may be one of the reasons for the relative scarcity of such studies focusing on developing countries (see e.g. [Laajaj & Macours \(2017\)](#) for a discussion, and [Villa \(2017\)](#); [Sánchez \(2017\)](#) for empirical applications). Both these studies have only looked at the relationship between parental investment and cognitive/non-cognitive skills (and/or health) but none of them have related the latter to work-related outcomes. Furthermore, our idea of using a multifaceted concept of work-related well-being poses additional empirical challenges due to the abstract nature of this concept which cannot be directly observed using a single measure.

Our paper is an attempt to address the above two challenges. First, we examine how this framework can be operationalised with a ‘limited’ data set (possibly a single cross-section) as it is often the only information set available in a developing country. Nevertheless, investigating the connections between educational investments and skills on the one hand, and skills and work-related wellbeing on the other, is extremely important for developing countries because these are fundamental ingredients of the human development process at the country and individual levels. For our study, we have chosen Bolivia, the poorest country in South America in monetary terms, in which the World Bank conducted a national survey on work related information at the individual level, which also includes some past information on the educational paths of these individuals. It is the 2012 Survey Towards Employability and Productivity (STEP). Although relying on a single cross-section prevents us from fully identifying certain dynamic aspects of [Cunha & Heckman \(2007\)](#)’s framework, we believe that our methodology does provide a way for identifying one of the key aspects of this framework, namely the role of skills in the link between education and work-related

wellbeing.

Second, drawing inspiration from the human development literature, we propose an outcome variable (work-related wellbeing) which is defined as a combination of multiple aspects of a job and hence not directly observable as such. Each of these aspects is specified as a latent variable and in turn measured by multiple indicators. The two are linked through measurement equations just as in the case of skills. [Cunha & Heckman \(2007\)](#) mention this possibility of having a *latent* outcome for anchoring skills (see page 898, section 3.5 of their paper) giving the example of a variable measured by a binary indicator, whereas in our case we have more than one latent outcomes with multiple indicators for each.

To sum up, in this paper we adapt the technology of skill formation framework to a context with limited available information and relate the skills to an unobservable (latent) concept that encompasses multiple aspects of work-related well-being.

The paper is structured as follows: in Section 2 we explain our theoretical model. In Section 3 we specify the functional forms used for the empirical implementation of our framework, describe the data used and state the identification conditions. The empirical results for Bolivia are discussed in Section 4. We end the paper with some concluding remarks in Section 5.

## 2 Theoretical Model

### 2.1 The conceptual framework

Our model structure can be depicted by the diagram in [Figure 1](#) representing a two-step process to understand the connection between educational investments and work-related well-being. The first step focuses on the effect of investments on skill acquisition, as per the technology of skill formation, accounting for individual heterogeneity and the unobservable nature of skills. In the second step, these skills expand people's

realistic chances to increase their work-related well-being as described in the previous section.

We consider work-related well-being as an intrinsically unobservable concept manifesting itself in multiple aspects measurable through observed indicators. We draw inspiration from one of the most influential contemporary approaches to human development, namely the Capability Approach ([Sen, 1980, 1985, 1999](#)), to propose this concept as our outcome of interest. According to this approach, human development is defined as the expansion of people’s choices in all dimensions so that they can achieve the lifestyle outcomes that they have reason to value. Applying this idea to the work dimension, we opt for a *vector* of interconnected work-related well-being aspects as the key outcome of interest.

[Figure 1 Here]

Skills are at the center of this framework; they mediate the relationship between educational investments and work-related well-being. In this study, we account for two important characteristics of skills that are well documented in this scholarship: they are themselves multifaceted and unobservable. Different types of skills are influenced by educational investments in different ways, and they generate different effects on labor market outcomes ([OECD, 2014](#); [Lindqvist & Vestman, 2011](#); [Heckman & Cunha, 2009](#)). Related studies regularly make the difference between *cognitive skills*, roughly defined as intelligence or acquired knowledge and *non-cognitive skills*, which refer to personality traits, behavior and socio-emotional abilities (see e.g. [Heckman et al. \(2006\)](#); [Prada & Urzúa \(2017\)](#)). Regarding the unobservable nature of skills, cognitive abilities manifest through some intelligence test scores or school grades, but these indicators are always partial accounts of this type of skill stock. Non-cognitive skills, in turn, can only be partially measured through some actions or reactions to psychological tests. As personality traits and behavior are impossible to observe perfectly, psychologists have developed different taxonomies for them ([Borghans et al., 2006](#); [Almlund et al., 2011](#)), but the most widely used one is the Big Five (see e.g. [Kautz](#)

et al. (2014); Bassi et al. (2012)). This taxonomy considers openness to experience, conscientiousness, extraversion, agreeableness and neuroticism as relevant elements of an individual's personality that are useful to better predict lifestyle outcomes in general (John & Srivastava, 1999; Duckworth et al., 2007), and labor market outcomes in particular (Heckman & Kautz, 2012, 2013). Many studies provide compelling evidence that both cognitive and non-cognitive skills are very important predictors of work-related well-being (Cunha & Heckman, 2007; Heckman & Cunha, 2009; Duncan & Magnuson, 2011; Hall & Farkas, 2011).

## 2.2 Model formulation

### 2.2.1 Structural equations

#### The process of skill formation

In the technology of skill formation, skill acquisition is represented as a dynamic process that evokes an overlapping generations model. Parents and family play a very powerful role in this process, especially during childhood and adolescence, not only because they endow genetics, but also because they define the child's environment in which the process takes place and decide if, when and how they make investments for the success of this process. Let us divide the timespan of skill acquisition of an individual into  $T$  development stages that we will denote as  $t = \{1 \dots T\}$ . The skill formation dynamics can be represented as follows:

$$\theta_{t+1} = g_{t+1}(\theta_t, I_t, \theta^p, u_{t+1}), \forall t \leq T \quad (1)$$

where  $\theta_{t+1}$  denotes the individual's stock of cognitive and non-cognitive skills at  $t + 1$  i.e.  $\theta_{t+1} = (\theta_{t+1}^C, \theta_{t+1}^{NC})'$ ,  $I_t$  denotes the investment made by the child's family (usually her parents) in period  $t$  to improve or expand skill stocks in the next period,  $t + 1$ . As time elapses and the child grows older, she gains autonomy with respect to educational

investments; thus  $I_t$  may also represent own efforts put into the skill formation process when  $t$  represents periods in time when the child is more mature and makes her own decisions.  $\theta^p$  represents the skill stock of the individual's parents, which is considered to be time-invariant.  $u_{t+1}$  represents shocks or other unobservable inputs in the process of skill formation. Allowing for generality, the skill formation function,  $g_{t+1}(\cdot)$  may vary in time. Equation (1) shows that the skills at any point in time depend on their stock in the previous periods: *together*, greater skills in period  $t$  induce greater stock of skills in the next period. This is called the *self-productivity* phenomenon (Heckman et al., 2006).

### Recursive substitution and the ‘final’ form of the dynamic system

Any attempt to empirically operationalize (1) and thereby uncover the rich dynamics that it contains requires longitudinal information on skill indicators. In the absence of such observations, for example in the context of a developing country, we propose to perform a recursive resolution of the skill equation by substituting the same equation in the previous period for past skills on the right hand side.

Successively substituting for past skills in (1), we obtain:

$$\theta_{t+1} = f_{1,t+1}(I_t, I_{t-1}, I_{t-2}, \dots, I_1, \theta^p, v_{1,t+1}) \quad (2)$$

where  $v_{1,t+1}$  is the resulting combined error term (including all errors  $u_{t+1}, u_t, u_{t-1}, \dots$ ) and  $f_{1,t+1}$  is appropriately defined. Notice that we have normalized the initial skill stocks to 1 without loss of generality, because skills do not have any specific unit of measurement being unobservable in nature, i.e. latent.

Equation (2) is usually called the ‘final’ form of the dynamic process. As past skills are eliminated from the equation, it does not allow us to identify the self-productivity phenomenon but it enables the framework to be applied, keeping all the other features, even in data deficient contexts. It only requires measurements of skills in one period ( $t + 1$ ), along with some information on previous investments on education. We argue

that the latter information may be easier to come by in cross-sectional surveys in the developing world, compared to richer and more detailed longitudinal observations on people’s skills.

Equation (2) allows us to capture the effects of investments on the different types of skills at different times. The magnitude of these effects depend on the extent to which skills are *malleable* by the schooling process. There is mixed evidence on this subject due to its obvious context-dependence (see e.g. [Carneiro et al. \(2007\)](#)). According to [Carneiro & Heckman \(2003\)](#), non-cognitive skills are often likely to be more malleable than cognitive skills, specially among young children. However, [Carlsson et al. \(2015\)](#) proved that some aspects of cognitive skills in adult populations can be effectively improved by short schooling periods. We propose a framework that is flexible enough to provide data-driven evidence about the malleable aspect of each type of skills, which is why we have the same general specification for each skill type.

Let us mention that transforming equation (1) into (2) is one plausible way to partially overcome data deficiencies in an attempt to uncover the skill formation process, as discussed in [Todd & Wolpin \(2003, 2007\)](#)<sup>3</sup>. They point out that in order to directly estimate equation (1), one would have to assume that either i) only contemporaneous investments in education are relevant for skill formation or ii) these investments are constant over time, and iii) that omitted investments and skill genetic endowments (or mental capacity) are uncorrelated with the observed investments. Note that including lagged investments in equation (2) relaxes i) and ii) which are indeed, severe, unrealistic assumptions. However, even equation (2) requires assumption iii) to hold, and there are many threats to the validity of this condition in our setting. For one, skills may shape investment decisions introducing reverse causality. Also, data deficiency forces us to leave out some relevant investments which will become part of the error term and may not be orthogonal to the observed investment variables. Finally, the observed investment indicators may themselves be measured with error. To effectively address all of these concerns, we will treat all our investment indicators as **endogenous** variables and add an explicit investment decision function which will provide us



with appropriate IVs for the investment indicators.

### Linking skills and work-related well-being

In order to connect skills and work-related well-being, let us recall that we divide the timespan of skill acquisition of an individual into  $T$  development stages. Work-related well-being is assessed in period  $T + 1$ , when the resources available to the individual include  $\theta_{T+1}$  i.e. the stock of cognitive and non-cognitive skills acquired up to that moment, and other exogenous resources and circumstances  $\tilde{x}_{T+1}$  that may influence her work-related well-being<sup>4</sup>. The relations describing the ‘conversion’ of resources into work-related well-being can be represented as follows:

$$Q_{T+1} = f_2(\theta_{T+1}, \tilde{x}_{T+1}, v_{2,T+1}) \quad (3)$$

where  $v_{2,T+1}$  is an error term.

### Possible endogeneity of educational investments

One crucial aspect highlighted in this literature concerns possible endogeneity of investments,  $I_t$  (see e.g. [Cunha et al. \(2010\)](#); [Heckman et al. \(2011\)](#)). This may arise from two sources: on the one hand, it may not be tenable to affirm that some shocks or unobservable inputs included in  $u_{1,t+1}$  (see eq. (1)) are not related to investments, whether these are made by the parents or by the individual herself. On the other hand, skills themselves may affect investments for their formation, inducing reverse causality with or without lags. Schooling is one example of such endogenous educational investments (see e.g., [Heckman et al. \(2006, 2011\)](#)).

[Cunha et al. \(2010\)](#) propose to divide the random shocks vector into two elements:  $(\pi_{t+1}, \epsilon_{t+1})$ , where  $\pi_{t+1}$  represents the unobservable inputs in the skill formation process that are correlated with or even partially identical to the unobservables in the investment decisions while being orthogonal to the other elements that enter the technology function  $g_{t+1}(\cdot)$ , and  $\epsilon_{t+1}$  a purely random shock independent of all explanatory

variables. Thus equation (2) may be rewritten as:

$$\theta_{t+1} = f_{1,t+1}(I_t, I_{t-1}, I_{t-2}, \dots, I_1, \theta^p, \pi_{t+1}, \epsilon_{t+1}), \forall t \leq T \quad (4)$$

To account for endogeneity of investments, an additional investment model or *investment policy function* is specified :

$$I_t = h_t(\theta_t, \theta^p, y_t, \pi_t), \forall t \leq T \quad (5)$$

where the presence of  $y_t$ , excluded from (4), should provide adequate instruments for  $I_t$ ,  $\forall t$ . As explained in Cunha et al. (2010) (see first line p.901 of this paper),  $y_t$  is made up of observable determinants of the investment decision, say family income, assets, constraints or other socioeconomic background variables, that do not directly affect skills.

As equation (1) explicitly accounts for an influence of  $\theta_t$  on  $I_t$ , in the absence of data on past skills (i.e. on  $\theta_t$ ), one can obtain a (partial) *reduced form* by replacing  $\theta_t$  by the skill production function (4) (for period  $t$ ) in (1) to obtain:

$$I_t = h_t(I_{t-1}, \dots, I_1, \theta^p, y_t, \tilde{v}_t) \quad (6)$$

with  $\tilde{v}_t$  being the new error term<sup>5</sup>.

### 2.2.2 Measurement equations for the latent variables

As we stressed before, skills are unobservable in nature (Cunha & Heckman, 2007; Heckman & Cunha, 2009; Kautz et al., 2014; Heckman et al., 2006). Cognitive skills partially manifest themselves through some intelligence test scores or school grades, which we will denote by  $Z^C$ . Non-cognitive abilities, in turn, may be partially measured through some actions or reactions to psychological tests, which we will denote by  $Z^{NC}$ . Thus regrouping all the observable indicators of both types of skills in period  $t + 1$

into a vector  $Z_{t+1}^S \equiv (Z_{t+1}^C, Z_{t+1}^{NC})'$ , we can write the skill measurement equations as:

$$Z_{t+1}^S = m_1(\theta_{t+1}, \varepsilon_{1,t+1}) \quad (7)$$

where  $\varepsilon_{1,t+1}$  represents unobservable elements that have an influence on skill indicators and  $m_1(\cdot)$  the transformation process of skills into observable indicators.

Similarly, as work-related well-being is a theoretical construct, it is hard to observe directly, but manifests itself in multiple ways. Hence we will assume  $Q_{T+1}$  to be latent measured by a set of observable job characteristics denoted as  $Z_{T+1}^Q$ . Thus we have the following measurement equations for work-related well-being:

$$Z_{T+1}^Q = m_2(Q_{T+1}, \varepsilon_{2,T+1}) \quad (8)$$

where  $m_2(\cdot)$  represents the process through which work-related well-being manifests itself in terms of observed job characteristics.  $\varepsilon_{2,T+1}$  represents a vector of shocks or unobservable factors in this process.

## 2.3 The full system

Regrouping all our equations into a system observed at time  $T + 1$  (adulthood), with retrospective information on educational investments in previous time periods, and some family background information at a single point in time in the past (during adolescence), we have:

$$\left\{ \begin{array}{lcl} \theta_{T+1} & = & f_{1,T+1}(I_T, I_{T-1}, I_{T-2}, \dots, I_1, \theta^p, v_{1,T+1}) \\ Q_{T+1} & = & f_2(\theta_{T+1}, \tilde{x}_{T+1}, v_{2,T+1}) \\ I_t & = & h_t(I_{t-1}, \dots, I_1, \theta^p, y_t, \tilde{v}_t), \text{ for } 1 \leq t \leq T; \text{ with } I_0 \text{ fixed} \\ Z_{T+1}^S & = & m_1(\theta_{T+1}, \varepsilon_{1,T+1}) \\ Z_{T+1}^Q & = & m_2(Q_{T+1}, \varepsilon_{2,T+1}) \end{array} \right. \quad (9)$$

System (9) is a Simultaneous Equation Model (SEM) with latent variables (Muthén, 1983, 1984). The first three vector equations represent the *structural* relations: the first one relates skill stocks (cognitive and non-cognitive) to educational investments, the second one relates different work-related well-being aspects (dimensions) to skills, and the third one investments to past investments, parental skills and other socio-economic characteristics of the household. The other two vector equations represent the *measurement* relations for the latent elements in our system, namely cognitive and non-cognitive skills, and the different dimensions of work-related well-being.

We argue that the three (vector) structural relations presented in (9) are important in their own right in the quest for understanding human development. Hence the functions  $f_{1,T+1}(\cdot)$ ,  $f_2(\cdot)$  and  $h_t(\cdot)$  should be estimated as such without deriving any reduced form and we go on to propose suitable means to estimate the full system *simultaneously*.

Before finishing this section, let us note that the standard Mincerian approach amounts to a simpler form of the structural equations which is obtained by substituting  $\theta_{T+1}$  in  $Q_{T+1}$ :

$$\begin{aligned} Q_{T+1} &= f_2(\theta_{T+1}, \tilde{x}_{T+1}, v_{2,T+1}) \\ &= f_2(f_{1,T+1}(I_T, I_{T-1}, \dots, I_1, \theta^p, v_{1,T+1}), \tilde{x}_{T+1}, v_{2,T+1}) \end{aligned} \tag{10}$$

This approach loses the richness of the explanation of labor market outcomes through the (dynamic) skill formation process.

## 3 Empirical strategy

### 3.1 The model

Let us now take the above model to the particular context of our empirical application on Bolivia. Based on the information available in our cross-sectional survey, we can

consider 3 periods in the lifespan of an individual for the structural equations: the early childhood (pre-school) period is taken to be the first period of skill formation  $t = 1$ ; the period of formal training in school covering primary to tertiary levels is  $t = 2$ . The subsequent period corresponds to the period of the survey for which we have observations on the individual's job and personal characteristics, and is therefore our  $T + 1$ . Thus we have  $T = 2$  (and  $T + 1 = 3$ ). We do not have any information on possible on-the-job training that the individuals may have acquired after finishing school, and our results have to be interpreted keeping this data limitation in mind.

Regarding investments, we take the 'school start gap' as the educational investment variable in period  $t = 1$ . This variable is measured as the age at which a child enters the schooling system (pre-school or primary) minus the normal age for starting compulsory primary schooling in Bolivia (6 years). So if children are enrolled in pre-school programmes, then the variable will be negative. On the other hand, if they start school later than 6, then it is positive<sup>6</sup>. Hence there is positive investment if the variable is negative. We argue that efforts to involve children in preschool activities play a positive role in the skill formation process ([Aizer & Cunha, 2012](#)). Conversely, postponing of the start of schooling might lead to delays in the skill acquisition process.

In period  $t = 2$  we consider years of education, i.e. time spent in formal education programs (primary, secondary and tertiary), as educational investments ([Trostel et al., 2002](#)). This variable represents efforts for skill acquisition through formal schooling.

The measurement equations use a) observable indicators of cognitive and non cognitive skills in adult life ( $Z_3^S$ ) and b) observable indicators of the job ( $Z_3^Q$ ). The exact content of these and other variables are discussed in section 3.3 ('Data and variable definitions').

Therefore, for our empirical setting, the system of equations (9) allowing for possible

endogeneity of the investments variables, becomes:

$$\left\{ \begin{array}{lcl} \theta_3 & = & f_1(I_1, I_2, \theta^p, v_{1,3}) \\ Q_3 & = & f_2(\theta_3, \tilde{x}_3, v_{2,3}) \\ I_2 & = & h_2(I_1, \theta^p, y_2, \tilde{v}_2) \\ I_1 & = & h_1(\theta^p, y_1, \tilde{v}_1) \\ Z_3^S & = & m_1(\theta_3, \varepsilon_{1,3}) \\ Z_3^Q & = & m_2(Q_3, \varepsilon_{2,3}) \end{array} \right. \quad (11)$$

These equations define a framework in which observable and unobservable elements are combined through structural and measurement relations. The observable elements are  $I_1, I_2, \theta^p, \tilde{x}_3, Z_3^S, Z_3^Q$ , as well as  $y_1$  and  $y_2$ . The unobservable elements are the skill stocks in adult life ( $\theta_3$ ), work-related well-being in adult life ( $Q_3$ ), as well as shocks captured in  $v_{1,3}, v_{2,3}, \tilde{v}_2, \tilde{v}_1$  and  $\varepsilon_{1,3}, \varepsilon_{2,3}$ . The vectors  $\theta_3, Q_3, I_2, I_1, Z_3^S, Z_3^Q$  are endogenous and  $\theta^p, \tilde{x}_3, y_2, y_1$  are assumed exogenous.

### 3.2 Identification and estimation procedure

Our identification strategy follows [Cunha et al. \(2010\)](#); [Cunha & Heckman \(2008\)](#); [Skrondal & Rabe-Hasketh \(2004\)](#); [Muthén \(1983, 1984\)](#); [Krishnakumar & Nagar \(2007\)](#) and it aims at securing consistent parameter estimates for the whole system, i.e. the structural equations as well as the measurement equations.

Let us start by discussing the identifying conditions of the structural equations. For notational convenience, let  $\eta = (\theta_3, Q_3, I_2, I_1)'$  be the vector of structural (latent and observed) variables containing, respectively, a 2-dimensional sub-vector of latent skills (cognitive,  $\theta_3^C$  and non-cognitive,  $\theta_3^{NC}$ ), an  $m$ -dimensional sub-vector of latent variables representing different dimensions of work-related well-being in period 3, i.e. current adult life, a  $p_2$ -dimensional vector of observed investment variables in period 2,  $I_2$  and a  $p_1$ -dimensional vector of observed investment variables in period 1,  $I_1$ . Thus  $\eta$  is sized  $(m + 2 + p_2 + p_1) \times 1$ .

Let us denote  $p_1 + p_2 = p$ . Finally, let  $X$  be a  $k \times 1$  vector containing exogenous control variables including parental characteristics  $\theta^P$ , current circumstances  $\tilde{x}_3$ ,  $y_2$ , and  $y_1$  i.e.  $X = (\theta^p, \tilde{x}_3, y_2, y_1)'$ . Denoting as  $k_1$  the number of elements in  $\tilde{x}_3$ ,  $k_2$  the number of elements in  $\theta_p$ ,  $\ell_1$  the number of elements in  $y_1$  and  $\ell_2$  the number of elements in  $y_2$ , we have  $k_1 + k_2 + \ell_1 + \ell_2 = k$ .

The first condition for identification consists in proposing additive linear forms, separable in  $v_{1,3}$ ,  $v_{2,3}$ ,  $\tilde{v}_2$  and  $\tilde{v}_1$  for functions  $f_1(\cdot)$ ,  $f_2(\cdot)$ ,  $h_2(\cdot)$ , and  $h_1(\cdot)$  respectively in (11)<sup>7</sup>, as is most commonly done in almost all empirical applications of this framework. Under these assumptions, the structural equations that relate the elements of  $\eta$  to each other as well as to the elements of  $X$  for the  $i$ -th individual can be expressed as:

$$A\eta - BX - v = 0 \quad (12)$$

where  $A$  and  $B$  are coefficient matrices and  $v$  is a vector containing the errors in the structural part of the system with  $E(v) = 0$  and full variance-covariance matrix  $V(v) = \Sigma$ .

The second set of identification conditions are the usual inclusion/exclusion restrictions in a simultaneous equation model that identify the coefficients of each structural equation. Regarding the identification of the equations for  $\theta_3$  and  $Q_3$ , we follow the logic of the technology of skill formation framework to impose these restrictions. We allow both types of skills to have a direct impact on work-related well-being  $Q_3$ , while educational investments are excluded from the equation for  $Q_3$ . These educational investments have a direct impact on both types of skills. Thus we explicitly posit that these investments only have an *indirect* effect on work-related well-being through skills. Parental skills are included in the (adult's) skill equations but not directly in the work-related wellbeing. Circumstances in current adult life are included in the equation for work-related well-being but not in the skill formation equations. Now, regarding the equations for  $I_2$  and  $I_1$ , it is easy to see that they are identified as long as there are enough instruments in  $y_1$  for  $I_1$  and in  $y_2$  for  $I_2$ . The mathematical trans-

lation of these theoretical restrictions imply the following configuration of elements in the structural equations (12):

$$\begin{bmatrix} \mathbf{I} & 0 & A_{11} & A_{12} \\ A_2 & \mathbf{I} & 0 & 0 \\ 0 & 0 & \mathbf{I} & A_3 \\ 0 & 0 & 0 & \mathbf{I} \end{bmatrix} \begin{bmatrix} \theta_3 \\ Q_3 \\ I_2 \\ I_1 \end{bmatrix} - \begin{bmatrix} B_1 & 0 & 0 & 0 \\ 0 & B_2 & 0 & 0 \\ 0 & 0 & B_3 & 0 \\ 0 & 0 & 0 & B_4 \end{bmatrix} \begin{bmatrix} \theta^p \\ \tilde{x}_3 \\ y_2 \\ y_1 \end{bmatrix} - \begin{bmatrix} v_{1,3} \\ v_{2,3} \\ \tilde{v}_2 \\ \tilde{v}_1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad (13)$$

where  $A_{11}$ ,  $A_{12}$ ,  $A_2$ ,  $A_3$ ,  $B_1$ ,  $B_2$ ,  $B_3$ , and  $B_4$  are corresponding coefficient matrices of appropriate dimensions and  $\mathbf{I}$  denotes the identity matrix.

Let us now turn to the identification conditions of the measurement equations in our system, i.e. the ones relating  $\theta_3$  and  $Q_3$  to their observable counterparts.

Here again, the first identifying condition consists of positing linear forms, separable in the latent variables and errors for vector equations  $m_1(\cdot)$  and  $m_2(\cdot)$  in (11). We assume that cognitive and non-cognitive skills are represented by a single latent factor each. But each one is associated with multiple indicators given by the vectors  $Z_3^C$  (for cognitive) and  $Z_3^{NC}$  (for non-cognitive). Work-related wellbeing is a latent vector of many aspects with multiple measurements for each aspect, given by the vector  $Z_3^Q$ . We will denote  $Z_3^S = (Z_3^C, Z_3^{NC})$ . Thus we have:

$$\begin{cases} Z_3^C &= \mu_C + \alpha_C \theta_3^C + \varepsilon_{1,3}^C \\ Z_3^{NC} &= \mu_{NC} + \alpha_{NC} \theta_3^{NC} + \varepsilon_{1,3}^{NC} \\ Z_3^Q &= \mu_Q + \alpha_Q Q_3 + \varepsilon_{1,3}^Q \end{cases} \quad (14)$$

where  $\alpha_j, j = \{C, NC, Q\}$  are vectors of factor loadings.

The second set of identifying conditions correspond to stochastic assumptions that are usual in this type of models (Skroendal & Rabe-Hasketh, 2004). We assume that vectors  $\varepsilon_{1,3}^j, j = \{C, NC, Q\}$  have zero mean and are independent two by two. But the measured indicators relating to a common latent variable will be correlated and this latent factor is the only channel for this correlation. We allow for the general case of



correlation between the two skills i.e.  $cov(\theta_3^C, \theta_3^{NC}) \neq 0$  in order to partially account for the complementary nature of these skills types in our static setting. We also allow for non-zero correlation between the different aspects (dimensions) of work-related well-being  $Q_3$  in order to capture common unmeasured surrounding circumstances that may influence the conversion process of well-being into actual measured outcome indicators. We further posit a scaling condition consisting of normalizing one element in each vector  $\alpha_j, j = \{C, NC, Q\}$  to unity.

Under these conditions, having at least three indicators for each latent variable secures identification of each equation in our measurement model (Muthén, 1983, 1984; Krishnakumar & Ballón, 2008). This condition is fulfilled for our skills and work-related wellbeing variables. At this point, let us mention that we do not specify measurement equations for investments as we only have one observed investment variable for each time-period. In this data deficient situation, it is not possible to treat investments as latent and include identifiable measurement equations to tackle possible measurement errors as done in Cunha et al. (2010). Instead, we assume that investments are directly observed and propose to deal with this consistency threat due to measurement errors by treating the observed investments as potentially endogenous variables and instrumenting them. Our results have to be interpreted keeping this in mind.

To estimate simultaneously the system formed by the structural equations (13) and measurement equations (14), we apply a maximum likelihood procedure as described in Muthén (1983, 1984). Let us collect the observed exogenous variables for the  $i$ -th individual in a vector  $\mathbf{G}_i = (y_1, y_2, \theta^P, \tilde{x}_3)'_i$  sized  $q_1 \times 1$ . Let us also collect the observed endogenous variables for this individual, including the observable indicators for the latent variables, in a vector  $\mathbf{F}_i = (Z_3^S, Z_3^Q, I_1, I_2)'_i$  sized  $q_2 \times 1$ . We will denote the number of observed variables in our system as  $q = q_1 + q_2$ , and the conditional moments of the observed endogenous variables as  $E(\mathbf{F}_i | \mathbf{G}_i) = \mu$  and  $V(\mathbf{F}_i | \mathbf{G}_i) = \Omega, \forall i$ . Note that the theoretical expressions of  $\mu$  and  $\Omega$  are derived after substituting the structural equations for  $\theta_3$  and  $Q_3$  into the measurement equations. Thus these expressions will contain all the parameters of the full system. Now, denote as  $\phi$  the

column vector of all the unknown parameters in our model, including the variance-covariance elements. Then, assuming conditional normality of  $F_i$  given  $G_i \forall i$ , the log-likelihood for each individual  $i$  can be written as:

$$\log L_i(\mathbf{F}_i | \mathbf{G}_i; \phi) = -\frac{1}{2}(q \times \log(2\pi) + \log(\det(\Omega)) + (\mathbf{F}_i - \mu)' \Omega^{-1} (\mathbf{F}_i - \mu))$$

and the log-likelihood of the system (omitting the arguments  $F_i$  and  $G_i \forall i$ ) is given by

$$\log L(\phi) = \sum_{i=1}^N \log L_i(\mathbf{F}_i | \mathbf{G}_i; \phi) \quad (15)$$

where  $N$  denotes the number of individuals in the sample. The above log-likelihood is maximised to obtain the parameter estimates  $\hat{\phi}$ . Robust (asymptotic) standard errors are computed using the heteroscedasticity-consistent sandwich formula for a Quasi-ML estimator.

Using these parameter estimates, scores for the latent variables  $\theta_3$  and  $Q_3$  can be calculated applying an empirical Bayes procedure. More details about this estimation technique can be found in [Muthén \(1983, 1984\)](#); [Skron dal & Rabe-Hasketh \(2004\)](#); [Krishnakumar & Ballón \(2008\)](#); [Krishnakumar & Nagar \(2007\)](#).

### 3.3 Data and variable definitions

We use data from the Bolivian household survey of the World Bank's Skills Towards Employability and Productivity (STEP) program, which was collected in 2012 and is publicly available through the World Bank's Microdata Catalog<sup>8</sup>. This survey is representative of the urban labor force in the country's main metropolitan areas, namely El Alto, La Paz, Cochabamba and Santa Cruz. We use this cross-sectional dataset because it is the only survey in the country that contains some suitable information for the construction of an empirical counterpart of the theoretical framework as we have described it.

As explained in the earlier sections, the static nature of our data does not allow us to fully grasp the dynamics behind the technology of skill formation. However, most of the other features of the framework, in particular the role of educational investments in the skill acquisition process and that of skills in determining work-related wellbeing can be captured as described in our theoretical sections. Our information on (past) family background variables and investments comes from the interviewees' responses to questions on the situation of their household when they were 15 years of age, and retrospective information about their schooling process. These are available in the same survey. It also has rich indicators of the stock of cognitive and non-cognitive skills in adulthood. Further, it has measures on different aspects of work-related well-being thus enabling us to capture its multifaceted nature to the maximum extent possible. Finally note that, as our past variables are only recorded for a single point in time in the past, we cannot apply measurement-corrective techniques to these variables.

Let us now explain the variables that we use in this study. Their descriptive statistics can be found in Appendix [A](#).

### **3.3.1 Measures of cognitive skills**

Our data set contains six observable indicators of cognitive skills resulting from written examinations.

The first indicator is called *Vocabulary*, which is calculated as the ratio between the number of correct responses in a 6-question written assessment and the time needed for completion of the test. This test was designed to determine whether the respondent is able to match written words of common everyday usage by adults with their corresponding figures. This variable is continuous and greater values indicate greater cognitive skills.

The second indicator is called *Sentence Processing* and it is calculated as the ratio between the number of correct responses in a 11-question written assessment and the time needed for completion of the test. In this test, respondents were asked to make

a dichotomous judgment about the sentence with respect to general knowledge about the state of world (true or false) or about the logic of the sentence (makes sense or not). This variable is also continuous and greater values indicate greater cognitive skills.

The third indicator is *Comprehensive Reading*, calculated as the ratio between the correct responses in a 17-question written assessment and the time required for the completion of the test. Greater values of this continuous variable depict greater cognitive skills. The test is comprised by three passages based on the kinds of text types that adults typically encounter. The respondents were required to choose the word that best completes a sentence in a passage.

The fourth indicator is called *Reading (self-reported)*, which relates to the size of the last document that the person recalls having read in the last 12 months, whether at work (if employed) or outside of work. This is a categorical variable in the 0-5 range depending on the maximum number of pages read; greater values of this variable depict greater cognitive skills.

The fifth indicator is called *Writing (self-reported)* is similar to the previous one, as it captures the number of pages (size) of the last document that the person recalls having written in the last twelve months whether at work or outside of work. It also ranges between 0 and 5 and greater values also denote greater cognitive skills.

The sixth indicator is called *Numeracy (self-reported)* and it adds up scores for various numerical calculations performed by a person whether at work or outside of work. Different scores have been assigned according to the complexity of these calculations: 0=none, 1=measuring sizes, weights or distances, 2=calculating prices or costs, 3=using fractions, decimals or percentages, 4=performing multiplications or divisions, 5=using advanced math. The variable takes up the maximum value between the scores obtained for numerical calculations at work or outside of work, thus ranging from 0 to 15. Greater values denote greater cognitive skills.

### 3.3.2 Measures of non-cognitive skills

Non-cognitive skills are measured by means of five personality traits in the Big Five taxonomy supplemented by three additional socio-emotional traits that have demonstrated a more adequate representation of a person's non-cognitive skills as *predictors* of life-style outcomes (see e.g. [Borghans et al. \(2006\)](#)). These socio-emotional traits are defined as grit, hostile attribution bias and decision-making.

The survey includes three questions for each personality trait or personality domain; each question is answered directly by the respondent using a four-point frequency scale going from *almost never* (1) to *almost always* (4). One summary measure of each personality trait is constructed as the simple average of the responses to the corresponding three questions.

Regarding the Big Five taxonomy, we will briefly define each domain and give an example of a question. Other questions can be found in the questionnaire available on the World Bank site given earlier. *Conscientiousness* is defined as the propensity to follow socially prescribed norms and rules, as well as to be goal-directed. One of the questions is *When doing a task, are you very careful?* *Openness to experience* is defined as enjoyment of learning and being confronted to new ideas. One of the questions is *Do you come up with ideas other people haven't thought of before?* *Neuroticism* refers to the tendency to feel negative emotions. One of the questions is *Do you worry a lot?* *Agreeableness* refers to the degree of orientation towards cooperation and empathy with other people. One of the questions is *Do you forgive other people easily?* Finally *extraversion* refers to the ability to be sociable and dominant while engaged in social interactions. One question is *Are you talkative?*

Likewise, the three additional socio-emotional traits are each grasped by three questions. The domain *grit* refers to perseverance for achieving long-term goals. One question is *Do you enjoy working on things that take a very long time (at least several months) to complete?* The domain *hostile attribution bias* refers the tendency to perceive hostile or mean intentions on the part of other people, even if, actually it is not

the case. One question is *Do people take advantage of you?* Finally, the domain of *decision-making* refers to the degree of thinking and consideration of multiple options when making important decisions that affect themselves or other people. One of the questions is *Do you think about how things you do will affect you in the future?*

### 3.3.3 Measures of work-related well-being

While the multifaceted nature of work-related well-being is now widely acknowledged in academic and political spheres (see e.g. [Anker et al. \(2008\)](#); [Ghai \(2008\)](#); [ILO \(2012\)](#); [OECD \(2014, 2015\)](#); [Lugo \(2007\)](#)), the precise evaluative dimensions and their quantitative representations are still far from the stage of reaching a worldwide consensus ([Vosko, 2002](#)). For instance ILO’s Decent Work agenda includes ten ‘substantive elements’ ([ILO, 2012](#)) or structural dimensions as follows: employment opportunities, adequate earnings, decent working time, combining work, family and personal life, work that should be abolished, stability and security of work, equal opportunity and treatment in employment, safe work environment, social security and social dialogue. Drawing upon the Sen-Stiglitz-Fitoussi Report on measurement of economic performance and social progress, similar job quality aspects are endorsed by the OECD ([OECD, 2014](#)). These are earnings quality, labor market security and quality of the work environment.

Although information on all the above aspects of employment would be ideal to include in our study, we are only able to obtain data on the following three dimensions of work-related wellbeing from our survey: i) employment opportunities and earnings, ii) decent working time and iii) safe work environment. Once again, in spite of this limitation, we go further with our exercise with the idea to make use of such data to draw meaningful conclusions.

STEP data provide us with multiple indicators of our first dimension, which is concerned with the availability of work positions and compensations, both monetary and non-monetary. We measure this dimension using three indicators: i) hourly earnings

(in logs) ii) contributions to social security and iii) firm size (number of workers). The rationale underpinning the latter indicator is drawn from vast evidence suggesting that smaller firms tend to trail larger ones in terms of stability, level and growth of wages as well as their capacity to offer formal positions in developing countries (see e.g. [Page & Soderbom \(2015\)](#); [Meghir et al. \(2015\)](#))).

Our second dimension (decent working time) has only one single indicator: hours worked per week beyond the hours established by law in the country, i.e. 48 hours for men and 40 hours for women. This indicator gauges the balance between working time and personal/leisure time and so we call this dimension *overtime work hours*. Higher values of this indicator imply a lower advantage for decent working time.

As we have only one indicator for this dimension, this is the only work-related well-being variable for which we cannot take account of measurement errors through a measurement equation. Hence it is supposed be a direct measure of the underlying latent variable.

Finally, our dataset also has information on the physical strain at work. According to health-related literature, physical intensity is an important risk factor at workplace (see e.g. [Widanarko et al. \(2015\)](#)). So our third dimension is concerned with the extent to which a job offers conditions that preserve and promote the physical integrity of a worker, among other things. We partially assess this dimension using three indicators: i) the extent to which the current job is perceived as physically demanding, ii) whether or not the job requires lifting items weighing above 50 pounds on a regular basis and iii) the worker's occupational category, ranging from 1= Soldier or unskilled worker to 10=Directive position in private or government offices.

### **3.3.4 Investments in skill formation and their instruments**

As explained in Section 3.1, in period  $t = 1$ , we take the school start gap (positive if delay in starting formal school and negative if enrolled in pre-school before 6) as the educational investment variable (there is positive investment if the variable is

negative). In period  $t = 2$  we consider years of education.

Drawing inspiration from [Cunha & Heckman \(2007, 2008\)](#); [Cunha et al. \(2010\)](#); [Trostel et al. \(2002\)](#), we consider socioeconomic and family background variables as possible instruments for investments in both periods 1 and 2. The only information available on past socioeconomic status of the individuals are the following: i) a 10-scaled variable grasping the perception of relative economic status when the respondent was 15 years of age (1=poorest; 10=richest), ii) the number of siblings in the household when the respondent was 15 years of age, and iii) the number of economic shocks experienced in respondent's household before she was 15, which may include robbery or bankruptcy.

Here we argue that past parental decisions on investments in education would have been influenced by the above past socioeconomic variables, but the latter may not have any direct influence on current skills or work-related wellbeing, after conditioning for parental skills. This idea finds support on substantial empirical evidence showing that parental skills configure many aspects of their children's lives in adulthood, including the family structure, social and cultural capital, learning environment, etc. (see e.g. [Tratamonte & Willms \(2010\)](#); [Björklund & Jäntti \(2009\)](#)). Let us point out that since we have three IVs for two endogenous investment variables, we are able to provide statistical proof of their validity by virtue of overidentification.

Note that as we only have these IVs at a particular point in time in the past, dynamic impacts of past skills are beyond the possible scope of this study, including determining whether parents invested in a compensatory or reinforcing manner in the past. However, these variables are adequate to implement an IV procedure for both the past investment variables. Also note that it is possible to have the same instrument set for both period investments as long as the remaining equations of the system contain enough exogenous variables that are excluded from the investment equations (which is the case in our model).



### 3.3.5 Control variables

Parental skills ( $\theta^p$ ) are gauged using a 4-scaled variable indicating the highest level of completed education of both mother and father according to the International Standard Classification of Education (ISCED, UNESCO). In this scale, 1=no formal education, 2=six years of primary education, 3=last three years of upper secondary education, and 4=post-secondary education or higher. Note that we cannot postulate a latent variable for parental skills as we only have one indicator for it. In other words we take years of education as a direct measurement of parental skills with no measurement error. An alternative interpretation could be that, instead of parental skills, we actually have parental years of education as the explanatory variable for skills.

Other exogenous resources and circumstances influencing the ability of having a good job ( $\tilde{x}_3$ ) include: i) a dummy variable indicating if the person works in a sector where work tends to use manual force (Agriculture/Fishing/Mining or Manufacturing/Construction) as opposed to other non-manual abilities (Commerce and other services), ii) age, iii) gender, and iv) a dummy variable indicating whether the individual has an indigenous language as a mother tongue - a proxy of her ethnic condition.

## 4 Results

### 4.1 The relevance of a multidimensional approach to work-related well-being

Let us begin by briefly making a case for the relevance of adopting a multidimensional approach to gauge work-related well-being by showing some preliminary evidence.

A one dimensional monetary viewpoint implies focusing on say earnings as the outcome of interest<sup>9</sup>. The results of this approach are given in panel (a) of Table 1. After correcting for endogeneity of investments, *only* cognitive skills should be considered as effective drivers of work-related well-being. Returns to non-cognitive skills are

insignificant. This would be a hasty bold claim as this one-dimensional framework does not allow us to adequately grasp the relative importance of both types of skills for work-related well-being. If we consider, for instance, contributions to social security as a one-dimensional indicator of the latter concept, results change considerably (see panel (b) of Table 1). Both cognitive and non-cognitive skills are effective drivers of work-related well-being as measured by this alternative indicator.

[Table 1 Here]

These results compellingly show that one-dimensional approaches to work-related well-being are insufficient to fully understand the mechanisms in operation.

## 4.2 The estimates of our full system

Before analyzing our results, let us present the main parameter estimates of system (11) formed by structural equations (13) and measurement equations (14).

Table 2 shows the main estimates of the structural equations, including the investment equations. Strictly following system (11), we have only included additional regressors in the work-related well-being equations. As a robustness test, we tried including them in the remaining equations as well but it does not result in any significant alteration.

Our results of the over-identification tests confirm that our IVs are valid thus justifying their exclusion in the skill equations (these results are reported in Appendix B).

Therefore the family background characteristics only have an indirect influence on skill stocks in adulthood through educational investments, after accounting for parental skills. This result is also linked with the strong evidence on inter-generational transmission of abilities (see e.g. Black et al. (2005), Heckman et al. (2006); Urzúa (2008)).

We now turn to the complementarity of the two types of skills. As mentioned before, even if we cannot account for the *dynamic* complementarity, our model does allow for *contemporaneous* complementarity between the two latent skill measures.

The coefficient of correlation between  $\theta_3^C$  and  $\theta_3^{NC}$ , which measures the extent of this complementarity, is estimated to be 0.11 and statistically significant at a 5% level.

We also estimate a variant of our system (11) which treats educational investments as exogenous and present its results in Appendix B (see Table (13)). A Hausman test comparing the two models rejects exogeneity of investments, which is why we retain the results presented in Table 2 for our analysis.

[Table 2 Here]

Let us now look at the parameter estimates for the measurement equations (Table 3). All indicators of cognitive skills load positively as expected and are highly significant. The ‘objective’ measures seem to have a higher loading than the self-reported ones. Similarly, all proposed indicators of non-cognitive skills load positively except for *Hostile behavior*. We also provide the goodness-of-fit statistic ( $R^2$ ) for each indicator. A better fit denotes a higher signal-noise ratio (see Appendix C).

[Table 3 Here]

### 4.3 The effects of skills

Our results confirm that cognitive and non-cognitive skills expand work-related well-being in general, although they have quite different effects over the three considered dimensions of this latter concept (Table 2). Cognitive skills expand all dimensions of work-related well-being; they exert the highest positive effect over employment opportunities and earnings, followed by safe work environment and finally, overtime work hours. In turn, non-cognitive skills exert the highest positive effect over safe work environment, followed by employment opportunities and earnings; they do not have a significant effect on overtime work hours. To delve deeper into these results, we derive standardised latent scores distributions of skills and the considered dimensions of well-being. Let us recall that standardised scores distributions have zero mean, thus positive valued scores depict individuals situated above average in the correspond-

ing distribution; the converse is true for negative valued scores. Furthermore, the standardised scores distributions have unit standard deviations, thus the score value represent the extent to which the considered individual deviates with respect to the ‘average’ individual in the distribution. Scores of employment opportunities and safe work environment increase with a bigger well-being, while higher scores of overtime work hours mean lower well-being levels. In all our interpretations, we will consider the *average* individual, for whom the score is zero, as the reference.

Table 4 shows the median values of well-being scores in each quintile of the distributions of both skill types. See Figures 2, 3 and 4 for a visual representation of the same values.

[Table 4 Here]

Let us start by focusing on our first dimension of work-related well-being, i.e. employment opportunities and earnings. Having higher levels of *both* types of skills has a positive effect over this well-being dimension. The median individual in the fifth quintile of the cognitive skills distribution has a well-being level that is 0.542 standard deviations higher than the reference. In contrast, the well-being of the median individual in the first quintile of the cognitive skills distribution is -1.016 standard deviations *lower* than the reference. The effect of non-cognitive skills on this dimension of work-related well-being is qualitatively similar albeit weaker in magnitude. The median individual in the fifth quintile of this type of skills has a well-being level that is 0.444 standard deviations higher than the reference; the corresponding figure is 0.636 lower for the median individual in the first quintile.

[Figure 2 Here]

Our results show that not only a lack of both cognitive and non-cognitive skills is associated with a relative disadvantage in terms of employment opportunities and earnings, but also that one has to be in the highest quintiles to have an above-average well-being, as the median individual in the fourth quintile is practically at the same level as the (overall) average individual for *both* types of skills. Enjoying high levels of well-being is an atypical situation for individuals sitting in the lower quintiles of both

skills distributions (see Figure 2). This is aligned with growing evidence of the importance of non-cognitive skills for employment and earnings in general (OECD, 2014; Heckman et al., 2006; Urzúa, 2008), and for Latin America in particular (Bassi et al., 2012). Also, it is interesting to visually notice that the higher the quintile, the higher dispersion of well-being levels. Sitting in the lower end of both skills distributions clearly puts Bolivian workers at a significant disadvantage in terms of this well-being dimension.

Let us now consider our second dimension of work-related well-being, namely overtime work hours (see Figure 3). Recall that the higher the value of overtime work hours the lower is the well-being level, and hence this indicator must be interpreted as an measure of *disadvantage*. We will only assess the effect of cognitive skills as we do not find evidence for significant effects of non-cognitive skills on this dimension.

[Figure 3 Here]

Contrary to the first dimension, the median person in every quintile, except for the first has a below-average well-being level in this dimension. The relative *disadvantage* in this dimension is concentrated among people with the *lowest* levels of cognitive skills; however, in absolute terms, this is not particularly strong. High well-being levels in this dimension are atypical across the *whole distribution* of cognitive skills, that is, overtime hours seem to occur at all levels of skills. It may be because overtime is only weakly associated with a lower quality job, and hence skills do not play a major role in it.

Similar evidence has been found across the region, including Honduras, Guatemala and el Salvador (Pagés & Madrigal, 2008), and some anecdotal evidence in Argentina (Bassi et al., 2012). This may also be the main reason why we do not detect statistically significant effects of non-cognitive skills on this particular dimension of work-related well-being.

Finally, let us consider our third dimension of work-related well-being, namely safe work environment (see Figure 4). Here, only the median individuals in the highest

cognitive skill quintiles (third to fifth) enjoy a well-being level that is higher than the reference. Similarly, in the case of non-cognitive skills, the median individuals from the third quintile upwards enjoy an above-average relative well-being advantage. Low well-being levels in this dimension are atypical only among the individuals sitting at the highest end of *both* types of skills. The effect of both types of skills is considerable and similar in magnitude. The median individual in the highest (fifth) quintile of cognitive skills has a well-being level that is 0.788 standard deviations higher than the reference and almost the same advantage (0.740 standard deviations higher) if it is the same quintile of non-cognitive skills. The median individual in the highest cognitive skills quintile has a well-being level that is 0.774 standard deviations higher than the reference and almost the same advantage (0.748 standard deviations higher) if it is the same quintile of non-cognitive skills. The similar importance of both types of skills for this particular dimension of work-related well-being is also a novel result for the country, but is aligned with the findings of [OECD \(2014\)](#). Both types of skills seem to enable workers to allocate their labor-force supply towards sectors and posts characterized by less physical strain, such as sales, trade or office-work (see e.g. [Bassi et al. \(2012\)](#)).

[Figure 4 Here]

Let us now discuss the effects of both types of skills relative to each other. On average, cognitive skills have greater positive effects compared to non-cognitive skills over all the considered dimensions of well-being. The mean effect of cognitive skills on employment opportunities and earnings, and safe work environment is around, respectively, 1.7 and 1.4 times that of non-cognitive skills (see Table 2). The superiority of cognitive skills over non-cognitive skills for the expansion of all the considered dimensions of well-being in Bolivia is in line with research conducted in other Latin American countries. For instance, [Díaz et al. \(2012\)](#) showed that the Peruvian labor market does not seem to give monetary rewards to non-cognitive skills such as agreeableness and cooperation, whereas it does reward basic cognitive skills similar to the ones that we measure here. Similarly, using the Colombian dataset of the World Bank’s STEP program, [Acosta et](#)

al. (2015) showed that cognitive skills are greatly associated with higher earnings and holding a formal job or a high-qualified occupation, while non-cognitive skills appear to have little direct influence on these labor market outcomes. It is interesting to note similar results in some OECD countries; referring to high-level cognitive skills, increases in this type of skills seem to outweigh increases in non-cognitive skills in terms of their contributions to better income and employment in Norway, Sweden and Switzerland, although it could be the contrary in other developed countries (OECD, 2014).

[Figure 5 Here]

One plausible explanation for this result in the Bolivian case lies in the fact that, in a 0-1 scale, the median worker has a latent factor score of 0.61 for non-cognitive skills and 0.32 for cognitive skills. Furthermore, as shown in Figure 5, the cumulative distribution of (normalised) non-cognitive skills stochastically dominates that of cognitive skills.

Thus the superiority of returns to cognitive skills relative to non-cognitive skills may be due to a relative scarcity of cognitive skills among workers. People with greater stocks of this type of skills may tend to be more valued in the job market. This result reflects the importance of improving basic cognitive skills for expanding work-related well-being in the country. Banerji et al. (2010) provide support for our statement, as based upon the 2006 PISA test for assessing quality of education, they show that 56% of students in lower middle-income countries have a *deficient* education in the sense that they face serious difficulty in effectively using mathematics to succeed in further education or work.

#### 4.4 The effects of educational investments on skill acquisition

This section discusses the estimation results of the structural equations describing the direct links between investments and skills. Let us recall that these equations spell out the channel through which educational investments have an effect on work-related

wellbeing, i.e. via skills.

From the estimated structural parameters reported in Table 2, one can see that years of schooling have significant direct influences on both cognitive and non-cognitive skills and through them, on work-related well-being. Greater levels of education induce greater levels of both types of skills and hence, expand this well-being. To have an idea about the magnitude of these influences, let us consider the individuals with average stocks of skills as the references.

As seen in Table 5, individuals with 12.5 years of schooling or higher tend to possess an above-average cognitive skill stock. In Bolivia, this corresponds to the completion of high school and short pre-college workshops. Shades of educational inequality stressed by Andersen (2001); UNDP (2010) are evident in light of this result, when we notice that only 26% of individuals in our data possess more than 12.5 years of education. Schooling above this level induce stocks of skills that may go up to being 1.547 standard deviations greater than the reference, which is the case of individuals having 23 years of schooling, a time period that would normally allow them to possess the highest tertiary academic degree. The median individual having finished undergraduate college studies, which normally takes 17 years of schooling in the country, has a stock of cognitive skills that is 0.705 standard deviations greater than the reference. Conversely, the median *uneducated* individual (zero years of schooling) has a stock of cognitive skills that is 1.983 standard deviations *lower* than the reference.

Concerning the effect of years of schooling over non-cognitive skills, a relative abundance of this type of skills is associated with people having more than 8 years of education, which corresponds to the time required for completion of primary education. Non-cognitive skills are distributed less unequally compared with cognitive skills, as 66% of individuals in our dataset have completed primary education. The median individual having the highest level of education (23 years) has a stock of non-cognitive skills that is 1.845 standard deviations greater than the reference. Conversely, the median uneducated individual has a stock of non-cognitive skills that is 1.450 lower than



the reference. The important magnitude of the effect on non-cognitive skills indicates that even if formal schooling programs tend to focus explicitly on the expansion of cognitive skills, they are, indirectly, important promoters of favourable personal traits through social interactions and involvement in situations that are useful to *forge* better attitudes and behavior. In that sense, our results provide novel empirical evidence for the Bolivian case that support similar findings in other contexts ([Kautz et al., 2014](#); [Bassi et al., 2012](#); [OECD, 2014](#); [Acosta et al., 2015](#)).

[Table 5 Here]

Finally, we find evidence for the fact that a late start of the schooling process is associated with lower stocks of cognitive skills (see Table 6). The median individual with 8 years of school-start gap has a stock of cognitive skills that is 1.436 standard deviations *lower* than the reference. Conversely, we find that avoiding school-start gaps and promoting an earlier insertion in schooling programs (i.e. negative school-start gaps) are beneficial situations for the acquisition of cognitive skills. The median individual with 2 years of pre-school has a stock of skills that is 0.175 standard deviations greater than the reference. The negative effect of school-start gaps over cognitive skills goes in line with [Kautz et al. \(2014\)](#); [Bassi et al. \(2012\)](#); [OECD \(2014\)](#); [Aizer & Cunha \(2012\)](#), who largely argue that children who ‘start behind, stay behind’, as it becomes increasingly difficult to compensate for negative environments as children grow older.

[Table 6 Here]

## 5 Concluding remarks

Understanding the nexus between education, skills and work-related well-being is fundamental in the quest to foster human development. Although rigorous studies are increasingly undertaken in developed countries to identify the connections involved, our knowledge of the same in the developing world is much less. This is largely due to the unavailability of appropriate longitudinal data in many developing countries. This

study is an attempt to show that it is still possible to uncover some important relationships with the limited information available. To this effect, we adapt the technology of skill formation of [Cunha & Heckman \(2007\)](#) to suit the nature of data available i.e. a single cross-section with some information on past investments. No doubt we are not able to uncover fully the rich dynamics of the original framework ([Cunha & Heckman \(2007, 2008\)](#); [Cunha et al. \(2010\)](#)) as we are constrained to rely on a recursive solution of the dynamic process to eliminate past skills for which data are missing.

We estimate a Simultaneous Equation Model with latent variables using our cross-sectional survey (STEP-Bolivia 2012) to obtain some key effects in the skill formation process. We also link the skills to a multifaceted outcome variable namely work-related well-being, for a better alignment of our objective with the theoretical underpinnings of contemporary development paradigms.

Among our most salient results, we find that both cognitive and non-cognitive skills have positive effects on work-related well-being in Bolivia, with the effect of the former being greater than that of the latter. Results in the literature are mixed concerning which type of skills is more important for a good job. To the best of our knowledge, this study provides the first empirical evidence on this debate for the Bolivian case.

Years of schooling are effective drivers of both cognitive and non-cognitive skills, but with different intensities. Relatively advantaged individuals with respect to their stock of cognitive skills are concentrated among people who have spent at least 12.5 years in the educational system. Relatively advantaged individuals in terms of their stock of non-cognitive skills have only completed primary education (8 years of schooling). Avoiding delays in the enrollment of children in school as well as exposing them to pre-school activities favor the development of cognitive skills.

Fostering multidimensional wellbeing is increasingly being recognised as a key objective of development and policy agendas throughout the world. Thus a proper understanding of the different pathways involved in the link between educational investments and work-related well-being is, among other things, crucial to the design of an adequate

policy. We believe that our study makes a small move towards drawing attention to the usefulness of such an analysis for an efficient action on the ground.

## References

- Acosta, P., Muller, N., Sarzosa, M. A. (2015). *Beyond qualifications: returns to cognitive and socio-emotional skills in Colombia*. World Bank Policy Research Working Paper, 7430.
- Almlund, M., Duckworth, A.L., Heckman, J., Kautz, T. (2011). *Personality psychology and economics*. In Hanushek, E.A, Machin, S., Woßmann, L. (Eds.) *Handbook of the Economics of Education* (volume 4, pp. 1-181). North Holland
- Andersen, L. (2001). *Low social mobility in Bolivia: Causes and consequences for development*. Working Paper 03/2001. Catholic University of Bolivia.
- Anker, R., Chernyshev, I., Egger, P., Mehran, F., Ritter J. (2008). Measuring decent work with statistical indicators. *International Labour Review*, 142(2), pp. 147-178. doi: [10.1111/j.1564-913X.2003.tb00257.x](https://doi.org/10.1111/j.1564-913X.2003.tb00257.x)
- Aizer, A., Cunha, F. (2012). *The production of human capital: Endowments, investments and fertility*. NBER Working Paper No. 18429. doi: [10.3386/w18429](https://doi.org/10.3386/w18429)
- Banerji, A., Cunningham, W., Fiszbein, A., King, E., Patrinos, H., Robalino, D., Tan, J. P. (2010). *Stepping up skills for more jobs and higher productivity*. World Bank, Washington, DC.
- Bassi, M., Busso, M., Urzúa, S., Vargas, J. (2012). *Disconnected: Skills, education, and employment in Latin America*. Inter-American Development Bank. Washington-DC, USA.
- Becker, G.S. (1964), *Human capital. A theoretical and empirical analysis with special reference to education* New York: Columbia University Press.
- Björklund, A., & Jäntti, M. (2009). Intergenerational income mobility and the role of family background. In: Salverda, W., Nolan, B., Smeeding, T.M. (Eds.), *Oxford Handbook of Economic Inequality*. Oxford University Press, Oxford, pp. 491-521.
- Black, S.E., Devereux, P.J., Salvanes, K.G. (2005). Why the apple doesn't fall far:

- Understanding intergenerational transmission of human capital. *American Economic Review*, 95(1), pp. 437-449. doi: [10.1257/0002828053828635](https://doi.org/10.1257/0002828053828635)
- Borghans, L., Lee, A., Heckman, J., Ter Weel, B. (2006). The economics and psychology of personality traits. *The Journal of Human Resources*, 43(4), pp. 972-1059. doi: [10.3368/jhr.43.4.972](https://doi.org/10.3368/jhr.43.4.972)
- Brunello, G., Schlotter, M. (2011). *Non cognitive skills and personality traits: Labour market relevance and their development in education & training systems*. IZA Discussion Papers No. 5743.
- Carlsson, M., Dahl, G.B., Ockert, B., Rooth, D.O. (2015). The effect of schooling on cognitive skills. *Review of Economics and Statistics*, 97(3), pp. 533-547. doi: [10.1162/REST\\_a\\_00501](https://doi.org/10.1162/REST_a_00501)
- Carneiro, P., Crawford, C., Goodman, A. (2007). *The impact of early cognitive and non-cognitive skills on later outcomes*. Centre for the Economics of Education, LSE, UK
- Carneiro, P.M., Heckman, J. (2003). *Human Capital Policy*. IZA Discussion Paper No. 821.
- Cunha, F., Heckman, J., Schennach, S. (2010). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica* 78, pp. 883-931. doi: [10.3982/ECTA6551](https://doi.org/10.3982/ECTA6551)
- Cunha, F., Heckman, J. (2008). Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation. *Journal of Human Resources*, 43, pp. 738-782. doi: [10.3368/jhr.43.4.738](https://doi.org/10.3368/jhr.43.4.738)
- Cunha, F., Heckman, J. (2007). The technology of skill formation. *American Economic Review*, 97(2), pp. 31-47. doi: [10.1257/aer.97.2.31](https://doi.org/10.1257/aer.97.2.31)
- Díaz, J.J., Arias, O., Tudela, D.V. (2012). *Does perseverance pay as much as being*

- smart? The returns to cognitive and non-cognitive skills in urban Peru*. Working Paper, World Bank, Washington, DC.
- Drobnic, S., Beham, B., Prag, P., Drobnic, S. (2010). Good job, good life? Working conditions and quality of life in Europe. *Social Indicators Research*, 99(2), pp. 205-225. doi: [10.1007/s11205-010-9586-7](https://doi.org/10.1007/s11205-010-9586-7)
- Duncan, G., Magnuson, K. (2011). The nature and impact of early achievement skills, attention skills, and behavior problem. In Duncan, G. & Murnane, R. (Eds.), *Wither Opportunity? Rising Inequality, Schools, and Children's Life Chances*, (pp. 47-71). New York: Russell Sage Foundation.
- Duckworth, A.L., Peterson, C., Matthews, M. D., Kelly, D. R. (2007). Grit: perseverance and passion for long-term goals. *Journal of Personality and Social Psychology*, 92(6), pp. 1087-1101. doi: [10.1037/0022-3514.92.6.1087](https://doi.org/10.1037/0022-3514.92.6.1087)
- Fasih, T. (2008). *Linking education policy to labor market outcomes*. The World Bank. doi: [10.1596/978-0-8213-7509-9](https://doi.org/10.1596/978-0-8213-7509-9)
- Ghai, D. (2008). Decent work: Concept and indicators. *International Labour Review*, 142(2), pp. 113-145. doi: [10.1111/j.1564-913X.2003.tb00256.x](https://doi.org/10.1111/j.1564-913X.2003.tb00256.x)
- Goldberg, J., Smith, J. (2008). The effects of education on labour market outcomes. In Fiske, E. & Ladd, H. (Eds.), *Handbook of Research in Education Finance and Policy* (pp. 688-708). Routledge.
- Hall, M., Farkas, G. (2011). Adolescent cognitive skills, attitudinal/behavioral traits and career wages. *Social Forces*, 89(4), pp. 1261-1285. doi: [10.1093/sf/89.4.1261](https://doi.org/10.1093/sf/89.4.1261)
- Harper, C., Marcus, R., Moore, K. (2003). Enduring poverty and the conditions of childhood: lifecourse and intergenerational poverty transmissions. *World Development*, 31(3), pp. 535-554. doi: [10.1016/S0305-750X\(03\)00010-X](https://doi.org/10.1016/S0305-750X(03)00010-X)
- Heckman, J., Humphries, J.E., Veramendi, G. (2018). Returns to education: The

- causal effects of education on earnings, health, and smoking. *Journal of Political Economy*, 126(1), pp. 197-246. doi: [10.1086/698760](https://doi.org/10.1086/698760)
- Heckman, J. Kautz, T (2013). *Fostering and Measuring Skills: Interventions That Improve Character and Cognition*. IZA Discussion Papers 7750, Institute for the Study of Labor (IZA).
- Heckman, J., Kautz, T. (2012). Hard evidence on soft skills. *Labour Economics*, 19(4), pp. 451-464. doi: [10.3386/w18121](https://doi.org/10.3386/w18121)
- Heckman, J., Humphries, J.E., Urzúa, S., Veramendi, G. (2011) *The effects of educational choices on labor market, health and social outcomes*. Working Paper No. 2011-002 Human Development Capital and Economic Opportunity Working Group, University of Chicago.
- Heckman, J., Stixrud, J., Urzúa, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics*, 24(3), pp. 411-482. doi: [10.3386/w12006](https://doi.org/10.3386/w12006)
- Heckman, J., Cunha, F. (2009). The economics and psychology of inequality and human development. *Journal of the European Economic Association*, 7 (2-3), pp. 320-364. doi: [10.1162/JEEA.2009.7.2-3.320](https://doi.org/10.1162/JEEA.2009.7.2-3.320)
- Heckman, J., Pagés, C. (2003). *Law and Employment: Lessons from Latin America and the Caribbean* NBER Working Paper No. 10129. doi: [10.3386/10129](https://doi.org/10.3386/10129)
- Heckman, J., Lochner, L., Todd, P. (2003). *Fifty Years of Mincer Earnings Regressions*. NBER Working Paper No. 9732. doi: [10.3386/w9732](https://doi.org/10.3386/w9732).
- ILO (2012). *Decent work indicators: Concepts and definitions*. ILO Manual, International Labor Office, Geneva.
- ILO (2013). *Panorama laboral 2013, América Latina y el Caribe*. OIT, Lima, Peru
- John, O., Srivastava, S.(1999). The Big Five Trait Taxonomy: History, Measurement

- and Theoretical Perspectives. In Pervin, L. & John, O. (Eds.), *Handbook of Personality: Theory and Research* (pp. 102-139). The Guilford Press, New York
- Kautz, T., Heckman, J., Diris, R., Ter Weel, B., Borghans, L. (2014). *Fostering and measuring skills: Improving cognitive and non-cognitive skills to promote lifetime success* OECD Education Working Papers 110, OECD Publishing. doi: [10.1787/5jxsr7vr78f7-en](https://doi.org/10.1787/5jxsr7vr78f7-en)
- Krishnakumar, J., Ballón, P. (2008). Estimating basic capabilities: A structural equation model applied to Bolivia. *World Development*, 36(6), pp. 992-1010. doi: [10.1016/j.worlddev.2007.10.006](https://doi.org/10.1016/j.worlddev.2007.10.006)
- Krishnakumar, J., Nagar, A. (2007). On exact statistical properties of multidimensional indices based on principal components, factor analysis, MIMIC and structural equation models. *Social Indicators Research*, 86(3), pp. 481-496. doi: [10.1007/s11205-007-9181-8](https://doi.org/10.1007/s11205-007-9181-8)
- Laaaj, R., Macours, K. (2017). *Measuring skills in developing countries*. The World Bank.
- Leßmann, O. (2012). Applying the capability approach empirically: An overview with special attention to labor. *Management Revue*, 23(2), pp. 98-118.
- Lin, D., Lutter, R., Ruhm, C. J. (2018). Cognitive performance and labour market outcomes. *Labour Economics*, 51, pp. 121-135. doi: [10.1016/j.labeco.2017.12.008](https://doi.org/10.1016/j.labeco.2017.12.008)
- Lindqvist, E., Vestman, R. (2011). The labor market returns to cognitive and noncognitive ability: Evidence from the Swedish enlistment. *American Economic Journal: Applied Economics*, 3, pp. 101-128. doi: [10.1257/app.3.1.101](https://doi.org/10.1257/app.3.1.101)
- Lugo, M. (2007). Employment: A proposal for internationally comparable indicators. *Oxford Development Studies*, 35(4), pp. 361-378. doi: [10.1080/13600810701701889](https://doi.org/10.1080/13600810701701889)
- Meghir, C., Narita, R., Robin, J.M. (2015). Wages and informality in developing



- countries. *American Economic Review*, 105(4). pp. 1509-1546. doi: [10.1257/aer.20121110](https://doi.org/10.1257/aer.20121110)
- Mincer, J. (1974). *Schooling, experience, and earnings*. New York: Columbia University Press.
- Muñoz de Bustillo, R., Fernández-Macías, E., Antón, J. I., Esteve, F. (2011). *Measuring more than money: The social economics of job quality*. Edward Elgar Publishing.
- Muthén, B. (1984). A general structural equation model with dichotomous, ordered categorical and continuous latent variable indicators. *Psychometrika*, 49(1), pp. 115-132. doi: [10.1007/BF02294210](https://doi.org/10.1007/BF02294210)
- Muthén, B. (1983). Latent variable structural equation modeling with categorical Data. *Journal of Econometrics*, 22(1-2), pp. 43-65. doi: [10.1016/0304-4076\(83\)90093-3](https://doi.org/10.1016/0304-4076(83)90093-3)
- OECD (2015), *Skills for social progress: The power of social and emotional skills*, OECD Skills Studies, OECD Publishing. doi: [10.1787/9789264226159-en](https://doi.org/10.1787/9789264226159-en)
- OECD (2014). How good is your job? Measuring and assessing job quality. In *OECD Employment Outlook 2014*, OECD Publishing, Paris
- Page, J., Soderbom, M. (2015). Is small beautiful? Small enterprise, aid and employment in Africa. *African Development Review*, 27(S1), pp. 44-55. doi: [10.1111/1467-8268.12138](https://doi.org/10.1111/1467-8268.12138)
- Prada, M. F., Urzúa, S. (2017). One size does not fit all: Multiple dimensions of ability, college attendance, and earnings. *Journal of Labor Economics*, 35(4), pp. 953-991. doi: [10.1086/692477](https://doi.org/10.1086/692477)
- Pagés, C., Madrigal, L. (2008). *Is Informality a good measure of job quality?: Evidence from job satisfaction data*. IDB Publications (Working Papers) 4122, Inter-American Development Bank.
- Sehnbruch, K. (2007). From the quantity to the quality of employment. In Alkire, S.,

- Comim, F. & Qizilbash, M. (Eds.), *The Capability Approach in Human Development: Concepts, Applications and Measurement*, Cambridge University Press.
- Sánchez, A. (2017). The structural relationship between early nutrition, cognitive skills and non-cognitive skills in four developing countries. *Economics & Human Biology*, 27, pp. 33-54. doi: [10.1016/j.ehb.2017.04.001](https://doi.org/10.1016/j.ehb.2017.04.001)
- Sen, A.K. (1999). *Development as Freedom*. Oxford: Oxford University Press.
- Sen, A.K. (1985). *Commodities and Capabilities*. Amsterdam: North Holland.
- Sen, A.K. (1980). Equality of What? In *The Tanner Lecture on Human Values, I*, pp. 197-220. Cambridge: Cambridge University Press.
- Skrondal, A., Rabe-Hesketh, S. (2004). *Generalized latent variable modelling: Multilevel, longitudinal and structural equation models*. Boca Raton, FL: Chapman & Hall/CRC.
- Todd, P., Wolpin, K. (2003). On the specification and estimation of the production function for cognitive achievement. *The Economic Journal*, 113(485), pp.F3-F33. doi: [10.1111/1468-0297.00097](https://doi.org/10.1111/1468-0297.00097)
- Todd, P., Wolpin, K. (2007). The production of cognitive achievement in children: Home, school, and racial test score gaps. *Journal of Human capital*, 1(1), pp 91-136. doi: [10.1086/526401](https://doi.org/10.1086/526401)
- Tramonte, L., & Willms, J. D. (2010). Cultural capital and its effects on education outcomes. *Economics of Education Review*, 29(2), pp. 200-213. doi: [10.1016/j.econedurev.2009.06.003](https://doi.org/10.1016/j.econedurev.2009.06.003)
- Trostel, P., Walker, I., Woolley, P. (2002). Estimates of the economic return to schooling for 28 countries. *Labour Economics*, 9(1), pp. 1-16. doi: [10.1016/S0927-5371\(01\)00052-5](https://doi.org/10.1016/S0927-5371(01)00052-5)
- UN (2015). *Transforming our World: The 2030 Agenda for Sustainable Development*. UN, New York.

- UNDP (2010). *Los cambios detrás del cambio: Desigualdades y movilidad social en Bolivia*. UNDP National HDR.
- UNDP (1990). *Human Development Report*. U.K.: Oxford University Press.
- Urzúa, S. (2008). Racial labor market traps: the role of abilities and schooling choices. *Journal of Human Resources*, 43(4), pp. 919-971. doi: [10.3368/jhr.43.4.919](https://doi.org/10.3368/jhr.43.4.919)
- Villa, K.M. (2017). Multidimensional human capital formation in a developing country: Health, cognition and locus of control in the Philippines. *Economics & Human Biology*, 27, pp. 184-197. doi: [10.1016/j.ehb.2017.06.003](https://doi.org/10.1016/j.ehb.2017.06.003)
- Vosko L.F. (2002). Decent work, the shifting role of the ILO and the struggle for global social justice. *Global Social Policy*, 2(1), pp. 19-46. doi: [10.1177/1468018102002001093](https://doi.org/10.1177/1468018102002001093)
- Widanarko, B., Legg, S., Devereux, J., Stevenson, M. (2015). Interaction between physical and psychosocial work risk factors for low back symptoms and its consequences amongst Indonesian coal mining workers. *Applied ergonomics*, 46, pp.158-167. doi: [10.1016/j.apergo.2014.07.016](https://doi.org/10.1016/j.apergo.2014.07.016)

# Notes

<sup>1</sup>See <http://www.oecdbetterlifeindex.org/>

<sup>2</sup>See [http://ec.europa.eu/environment/beyond\\_gdp/download/bgdp-summary-notes.pdf](http://ec.europa.eu/environment/beyond_gdp/download/bgdp-summary-notes.pdf)

<sup>3</sup>Our equation (2) is similar to what they call the ‘cumulative specification’.

<sup>4</sup>Here we take the notion of resources in a large sense to include personal resources such as personal abilities and social capital, as well as institutional and infrastructural support.

<sup>5</sup>Note that for  $t = 1$ , the equation for  $I_1$  would have  $I_0$  on the right hand side; in the absence of any information on it we assume that it is fixed and constant for all.

<sup>6</sup>If they start school at the usual age of 6 without any pre-school, then it is zero.

<sup>7</sup>Note that the linearity assumption only concerns the parameters and includes the possibility of entering the variables in a transformed form (e.g. logged) in the equation. We will show in our empirical application that this assumption of linearity in parameters does not imply linear effects.

<sup>8</sup>website: [data.worldbank.org](http://data.worldbank.org)

<sup>9</sup>This is in fact a simplified version of our proposed framework in which  $Q_3$  has one dimension fully measured by one indicator:  $Z_3^Q = Q_3 = income$

# Tables

Table 1: Selected estimation results (non-standardized) of one-dimensional approaches to work-related well-being

	Work-related well-being measured by:					
	(a) Hourly earnings (log)			(b) Contributions to social security		
	W	C	NC	W	C	NC
Cognitive skills	0.101 **			0.155 ***		
Non-cognitive skills	-0.025			0.136***		
Years of schooling		0.467 ***	0.344 ***		0.467 ***	0.348 ***
Age school-start gap		-0.089 ***	-0.023		-0.089 ***	-0.024
Control variables	yes	yes	yes	yes	yes	yes
$R^2$	0.02	0.54	0.18	0.11	0.54	0.19
Obs. (full info.)		1,484			1,484	

W: Work-related well-being; C: Cognitive skills; NC: Non-cognitive skills

Robust standard errors. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.10$

Table 2: Estimated structural parameters (standardized)

	Dependent Variable						
	$Q_3$			$\theta_3$		$I$	
	$W_1$	$W_2$	$W_3$	C	NC	$I_2$	$I_1$
Cognitive skills (C)	0.243***	-0.133***	0.193***				
Non-cognitive skills (NC)	0.131**	0.027	0.126**				
Years of schooling ( $I_2$ )				0.468***	0.350***		
School start gap ( $I_1$ )				-0.089***	-0.023	-0.147	
Mother's education							
- ISCED <1 (ref.)							
- ISCED = 1				-0.006	-0.068*	0.061***	0.017
- ISCED = 2 or 3				0.135***	-0.015	0.056*	-0.067**
- ISCED >3				0.164***	-0.042	0.126***	-0.043
Father's education							
- ISCED <1 (ref.)							
- ISCED = 1				-0.042**	0.067**	0.117***	-0.008
- ISCED = 2 or 3				0.140***	0.080*	0.181***	-0.066**
- ISCED >3				0.196***	0.121**	0.292***	-0.118***
Socioeconomic status at age 14						0.094***	-0.053*
Brothers and sisters at age 14						-0.049	0.072**
Economic shocks before age 14						-0.115***	0.026
Indigenous (1=yes)	-0.045	0.036	-0.118***				
Works in manual post (1=yes)	-0.047*	-0.060**	-.156***				
Gender (1=Male)	-0.099***	0.011	0.171***				
Age	0.028	0.069***	-0.018				
Constant		0.570***				2.361***	0.034***
$R^2$	0.17	0.02	0.18	0.54	0.20	0.30	0.06
Obs. (full info.)				1,484			

$W_1$ : Employment opportunities and earnings;  $W_2$ : Overtime work hours;  $W_3$ : Safe work environment

ISCED: International standard classification of education

Robust standard errors. \*\*\*:  $p < 0.01$ , \*\*:  $p < 0.05$ , \*:  $p < 0.10$

Table 3: Estimations of Measurement Equations (standardized factor loadings)

Indicator	Associated Latent Variable				
$Z_3^C/Z_3^{NC}/Z_3^Q$	$W_1$	$W_3$	C	NC	$R^2$
Vocabulary			0.669***		0.45
Sentence processing			0.859***		0.74
Comprehensive reading			0.884***		0.78
Reading (self-reported)			0.473***		0.22
Writing (self-reported)			0.570***		0.32
Numeracy (self-reported)			0.420**		0.18
Grit				0.322***	0.10
Decision making				0.351***	0.12
Hostile behavior				-0.149***	0.02
Extraversion				0.422***	0.18
Conscientiousness				0.340***	0.12
Openness to experience				0.549***	0.30
Emotional stability				0.183***	0.03
Agreeableness				0.272***	0.07
Hourly earnings	0.125***				0.02
Firm size	0.770***				0.59
Contributions to social security	0.791***				0.63
Physical demand at work		0.688***			0.47
Not heavy weight lift at work		0.755***			0.57
Occupational category		0.339***			0.12
Obs. (full info.)			1,484		

$W_1$ : Employment opportunities and earnings

$W_3$ : Safe work environment; C: Cognitive skills; NC: Non-cognitive skills

\*\*\*: p-value < 0.01

Table 4: Median standardised latent scores of work-related well-being by quintile of skills

Quintile	Cognitive skills			Non-cognitive skills		
	$W_1$	$W_2$	$W_3$	$W_1$	$W_2^{(+)}$	$W_3$
1 (lowest)	-1.016	0.033	-0.497	-0.636		-0.413
2	-0.357	-0.333	-0.281	-0.331		-0.061
3	-0.158	-0.333	0.223	-0.148		0.267
4	0.038	-0.578	0.450	-0.040		0.354
5 (highest)	0.542	-0.364	0.788	0.444		0.740

(+): Not statistically significant relation

$W_1$ : Employment opportunities and earnings

$W_2$ : Overtime work hours

$W_3$ : Safe work environment

Table 5: Median standardized latent scores of skills by years of schooling

Years of schooling	Cognitive skills	Non-cognitive skills
0	-1.983	-1.450
1	-1.745	-1.756
2	-1.665	-0.967
3	-1.543	-1.171
4	-1.421	-0.771
5	-1.207	-0.771
6	-1.014	-0.759
8	-0.351	-0.382
11	-0.002	0.157
12	-0.101	0.020
12.5	0.344	0.157
15	0.326	0.450
16	0.330	0.419
17	0.705	0.744
19	1.092	1.089
23	1.547	1.845

Table 6: Median standardized latent scores of cognitive skills by school-start gap

School-start gap	Cognitive skills
-2	0.175
-1	0.254
0	0.161
1	-0.436
2	-1.160
3	-1.233
4	-1.535
5	-1.493
6	-1.704
8	-1.463



## Figures

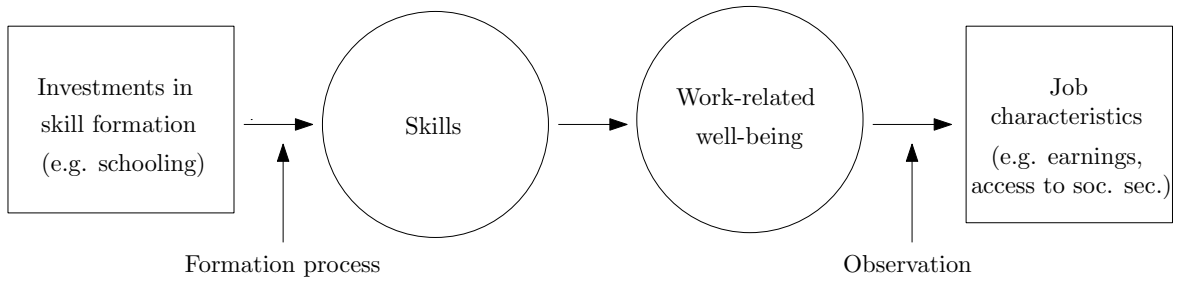


Figure 1: Analytical framework in a diagram

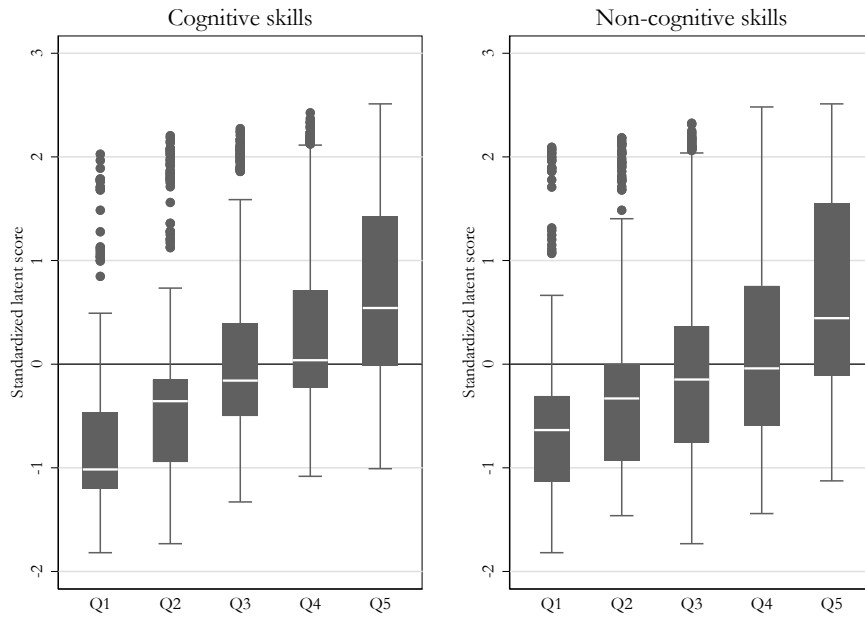


Figure 2: Standardised latent scores of  $W_1$  (employment opportunities and earnings) by skill quintiles

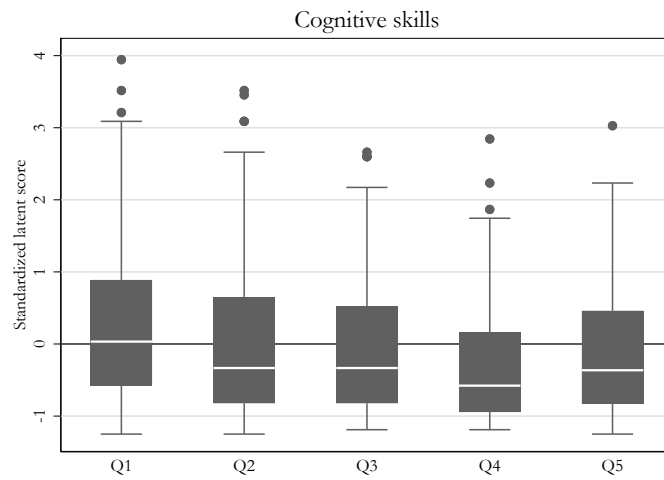


Figure 3: Standardised latent scores of Overtime Work Hours by quintiles of cognitive skills

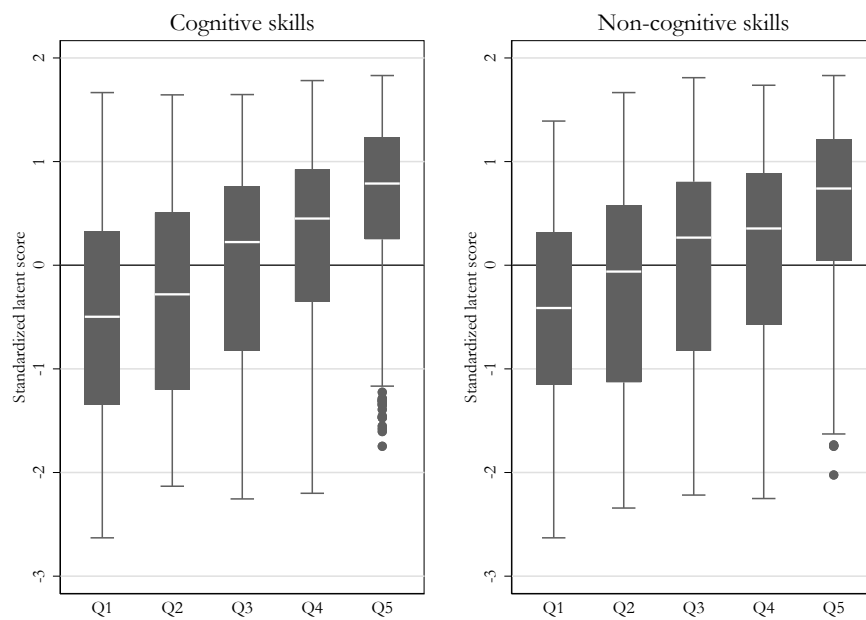


Figure 4: Standardised latent scores of Safe Work Environment by quintiles of cognitive skills

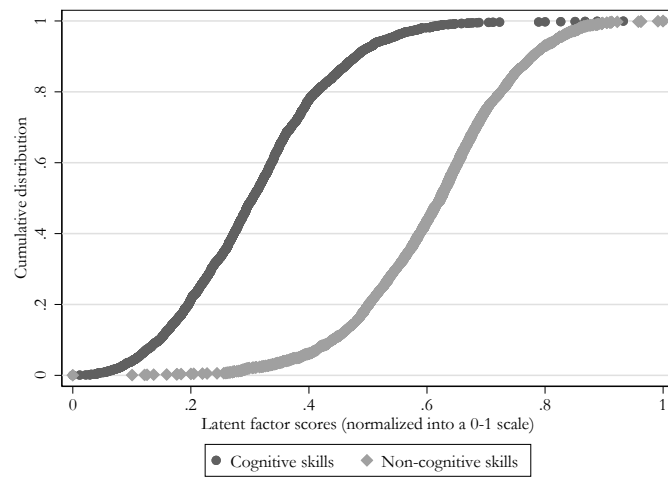


Figure 5: Cumulative distribution of skills

# Appendix

## A Descriptive statistics

Table 7: Cognitive skills indicators: summary statistics (non-standardized)

Indicator	Mean	St. Dev.	Min.	Max.
Vocabulary	0.038	0.023	0	0.545
Sentence processing	0.044	0.025	0	0.250
Comprehensive reading	0.033	0.019	0	0.157
Reading (self-reported)	3.232	1.744	0	5
Writing (self-reported)	1.948	1.830	0	5
Numeracy (self-reported)	9.490	4.056	0	15

Source: STEP Bolivia, 2012

Table 8: Non-cognitive skills indicators: summary statistics (non-standardized)

Indicator	Mean	St. Dev.	Min.	Max.
Conscientiousness	3.125	0.530	1	4
Openness to experience	3.181	0.566	1	4
Neuroticism	2.431	0.702	1	4
Agreeableness	3.044	0.643	1	4
Extraversion	2.988	0.706	1	4
Grit	2.936	0.641	1	4
Hostile attribution bias	3.028	0.625	1	4
Decision-making	1.892	0.687	1	4

Source: STEP Bolivia, 2012

Table 9: Work-related well-being variables: summary statistics

Dimension	Indicator	Mean	St. Dev.	Min.	Max.
Employment opportunities and earnings	Hourly earnings (log, 2011 USD)	2.184	2.142	-50.657	7.676
	Contributions to social security	0.243	0.429	0	1
	Firm size	1.804	0.756	1	3
	Overtime work hours	6.297	13.191	0	86
Safe work environment	Physical demand	6.132	2.388	1	10
	Lifting weights	0.565	0.496	0	1
	Occupational category	4.857	2.457	1	10

Source: STEP Bolivia, 2012

Table 10: Education investments: summary statistics

Variable	Mean	St. Dev.	Min.	Max.
School start gap	0.185	0.918	-2	8
Years of education	11.204	4.380	0	23

Source: STEP Bolivia, 2012

Table 11: IVs: summary statistics

Variable	Mean	St. Dev.	Min.	Max.
Socioeconomic status at age 15	4.450	1.745	1	10
Number of siblings in hh. at age 15	3.768	2.368	0	19
Number of economic shocks in hh. at age 15	1.369	1.638	0	10

Source: STEP Bolivia, 2012

Table 12: Control variables: summary statistics

Variable	Mean/Freq.	St. Dev.	Min.	Max.
Indigenous (1=yes)	0.260	0.439	0	1
Manual worker (1=yes)	0.272	0.445	0	1
Gender (1=Male)	0.577	0.494	0	1
Age	33.065	13.129	15	64
Mother's education				
- None	48.26			
- Primary	8.51			
- Secondary	25.61			
- Tertiary	17.62			
Father's education				
- None	34.34			
- Primary	9.19			
- Secondary	29.97			
- Tertiary	26.51			

Source: STEP Bolivia, 2012

## B Exogenous investments and IV validity

### Exogenous investments

Table 13: Standardized parameters

	Dependent Variable				
	$Q_3$			$\theta_3$	
	$W_1$	$W_2$	$W_3$	C	NC
Cognitive skills (C)	0.304***	-0.122***	0.205***		
Non-cognitive skills (NC)	0.192***	0.036	0.146**		
Years of schooling ( $I_2$ )				0.476***	0.384*
School start gap ( $I_1$ )				-0.087***	-0.023
Mother's education					
- ISCED <1 (ref.)					
- ISCED = 1				-0.006	-0.067*
- ISCED = 2 or 3				0.134***	-0.015
- ISCED >3				0.162***	-0.041
Father's education					
- ISCED <1 (ref.)					
- ISCED = 1				-0.041**	0.064**
- ISCED = 2 or 3				0.138***	0.073*
- ISCED >3				0.196***	0.108**
Socioeconomic status at age 14					
Brothers and sisters at age 14					
Economic shocks before age 14					
Indigenous (1=yes)	-0.040	0.036	-0.117***		
Works in manual post (1=yes)	-0.053*	-0.059**	-.160***		
Gender (1=Male)	-0.101***	0.009	0.169***		
Age	0.028	0.062**	-0.013		
Constant		0.562***			
$R^2$	0.21	0.02	0.19	0.55	0.22

$W_1$ : Employment opportunities and earnings;

$W_2$ : Overtime work hours;  $W_3$ : Safe work environment

ISCED: International standard classification of education

Robust standard errors. \*\*\*:  $p < 0.01$ , \*\*:  $p < 0.05$ , \*:  $p < 0.10$

A Hausman test comparing these estimates (and the elements of the variance-covariance matrices) with those of the model with endogenous investments confirms that the former are inconsistent:  $\chi^2 = 95.469$ ;  $DF = 29$ ;  $p$ -value: 0.000

### IV validity

Sargan overidentification tests are implemented to assess the validity of our exclusion

restrictions. We use the latent factor scores of skills (from a factor analysis of the skill measurement equations to run IV regressions of skills on investments and parental skills, and implement an overidentification test on the instrument set. Note that these latent scores are used in place of the true scores that are unobserved; however as they are the explained variable in this case, the difference between the true and estimated scores becomes part of the error term of the equations. As seen below, the null hypothesis is accepted for both skill equations, thus confirming the validity of the instruments used for our investment variables.

Table 14: Sargan overidentification tests

Endogenous explained	Endogenous explanatory	Exogenous included	Exogenous excluded (instruments)	p-value
$\hat{\theta}_3^C$	$I_1, I_2$	$\theta^P$	$y_1, y_2$	0.761
$\hat{\theta}_3^{NC}$	$I_1, I_2$	$\theta^P$	$y_1, y_2$	0.313

## C Signal to noise ratios

Table 15: Signal and noise in measurement variables

Associated latent variable	Measurement	Signal(%)	Noise(%)
<i>C</i>	Vocabulary	0.45	0.55
<i>C</i>	Sentence processing	0.74	0.26
<i>C</i>	Comprehensive reading	0.78	0.22
<i>C</i>	Reading (self-reported)	0.22	0.78
<i>C</i>	Writing (self-reported)	0.32	0.68
<i>C</i>	Numeracy (self-reported)	0.18	0.82
<i>NC</i>	Grit	0.10	0.90
<i>NC</i>	Decision making	0.12	0.88
<i>NC</i>	Hostile behavior	0.02	0.98
<i>NC</i>	Extraversion	0.18	0.82
<i>NC</i>	Conscientiousness	0.12	0.88
<i>NC</i>	Openness to experience	0.30	0.70
<i>NC</i>	Emotional stability	0.03	0.97
<i>NC</i>	Agreeableness	0.07	0.93
<i>W</i> <sub>1</sub>	Hourly earnings	0.02	0.98
<i>W</i> <sub>1</sub>	Firm size	0.59	0.41
<i>W</i> <sub>1</sub>	Contributions to social security	0.63	0.37
<i>W</i> <sub>3</sub>	Physical demand at work	0.47	0.53
<i>W</i> <sub>3</sub>	Not heavy weight lift at work	0.57	0.43
<i>W</i> <sub>3</sub>	Occupational category	0.12	0.88

*W*<sub>1</sub>: Employment opportunities and earnings

*W*<sub>3</sub>: Safe work environment; *C*: Cognitive skills; *NC*: Non-cognitive skills