

Wp

WORKING PAPERS

EDUCATION, SIGNALING AND THE ALLOCATION OF ENTREPRENEURIAL SKILLS
10/2017
N° 2017/15

EDUCATION, SIGNALING AND THE ALLOCATION OF ENTREPRENEURIAL SKILLS

Arozamena, L.
Ruffo, H.

EDUCATION, SIGNALING AND THE ALLOCATION OF ENTREPRENEURIAL SKILLS

Arozamena, L.
Ruffo, H.

CAF – Working paper N° 2017/15
10/2017

ABSTRACT

We assess the allocative importance of education when workers can choose to self-employ. To do so, we build a model combining educational choices with the labor market and self-employment. Education can increase workers' human capital and may signal their ability as well. Both roles can be more important for working in a firm than for self-employment. We show that when education performs worse its signaling role, firms cannot distinguish high and low productivity workers, and there is a higher proportion of workers that allocate in less productive activities as self-employed. This option further reduces incentives to educate, given that education is less valuable for a worker if self-employed. Lowering the cost of education increases the number of educated workers, but does not solve the signaling problem, and could generate stronger misallocation.

Small sections of text, that are less than two paragraphs, may be quoted without explicit permission as long as this document is stated. Findings, interpretations and conclusions expressed in this publication are the sole responsibility of its author(s), and it cannot be, in any way, attributed to CAF, its Executive Directors or the countries they represent. CAF does not guarantee the accuracy of the data included in this publication and is not, in any way, responsible for any consequences resulting from its use.

© 2017 Corporación Andina de Fomento

EDUCACIÓN, SEÑALIZACIÓN Y ASIGNACIÓN DE HABILIDADES EMPRESARIALES

Arozamena, L.
Ruffo, H.

CAF - Documento de trabajo N° 2017/15
10/2017

RESUMEN

Evaluamos la importancia de la educación como mecanismo de asignación cuando los trabajadores pueden elegir autoemplearse. Para ello, construimos un modelo que combina elecciones educativas, un mercado laboral y el autoempleo. La educación puede aumentar el capital humano de los trabajadores y también puede ser una señal de su capacidad. Ambos roles pueden ser más importantes para trabajar en una empresa que para el autoempleo. Demostramos que cuando la educación desempeña peor su papel de señalización, las empresas no pueden distinguir a los trabajadores de alta y baja productividad, y existe una mayor proporción de trabajadores que se dedican a actividades menos productivas como trabajadores autónomos. Esta opción reduce aún más los incentivos para formarse, dado que la educación es menos valiosa para un trabajador por cuenta propia. La reducción del costo de la educación aumenta el número de trabajadores educados, pero no resuelve el problema de la señalización, y podría generar una peor asignación.

Small sections of text, that are less than two paragraphs, may be quoted without explicit permission as long as this document is stated. Findings, interpretations and conclusions expressed in this publication are the sole responsibility of its author(s), and it cannot be, in any way, attributed to CAF, its Executive Directors or the countries they represent. CAF does not guarantee the accuracy of the data included in this publication and is not, in any way, responsible for any consequences resulting from its use.

© 2017 Corporación Andina de Fomento

Education, signaling and the allocation of entrepreneurial skills

Leandro Arozamena and Hernán Ruffo¹
Universidad Torcuato Di Tella

October 2017

Abstract We assess the allocative importance of education when workers can choose to self-employ. To do so, we build a model combining educational choices with the labor market and self-employment. Education can increase workers' human capital and may signal their ability as well. Both roles can be more important for working in a firm than for self-employment. We show that when education performs worse its signaling role, firms cannot distinguish high and low productivity workers, and there is a higher proportion of workers that allocate in less productive activities as self-employed. This option further reduces incentives to educate, given that education is less valuable for a worker if self-employed. Lowering the cost of education increases the number of educated workers, but does not solve the signaling problem, and could generate stronger misallocation.

1 Introduction

Entrepreneurial talent misallocation is one of the main driving forces behind productivity differences across countries (Hopenhayn, 2014). This problem arises when such abilities are not fully exploited within firms. As a consequence, fewer or lower-quality firms are created, which results in lower average productivity. How entrepreneurial talent is allocated in the labor market relates to the options available and the constraints faced by workers when making their occupational choices.

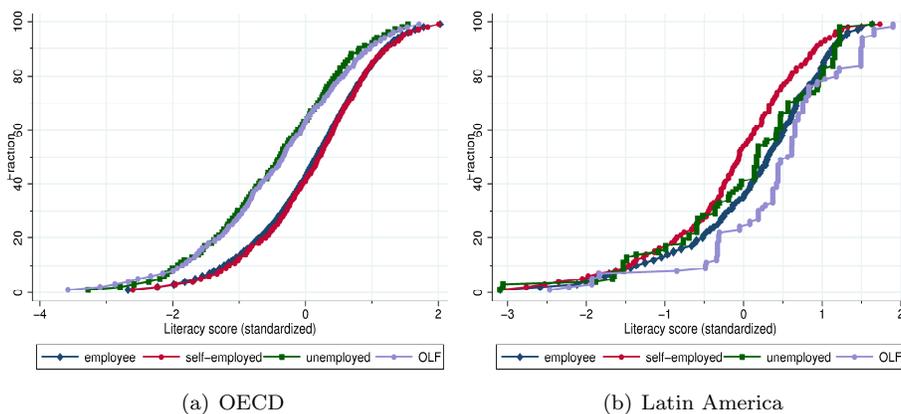
The interaction between entry costs and financial frictions is frequently mentioned as one of the main reasons why entrepreneurial talent is not allocated optimally (Banerjee and Newman, 1993, Buera et al., 2011). These obstacles prevent the creation of firms and distort the decisions of truly able entrepreneurs, thus negatively affecting not only the quantity of new firms, but also their quality. Take, for example, financial constraints. If wealth is in the hands of less able entrepreneurs, then the quality of new firms will be lower (Erosa and Allub, 2012, Banerjee and Newman, 1993). Additionally, new firms could be oriented to less capital-intensive and less productive sectors, as in Buera et al. (2011), or continue operating informally because of the higher costs of entry in the formal sector (D'Erasmus and Moscoso Boedo, 2012). Survival rates could also be lower because of these barriers (Midrigan and Xu, 2014). Young firms are typically credit constrained, while reduced information and managerial abilities could also lower their survival probability (Mano et al., 2012).

¹E-mail: larozamena@utdt.edu, hruffo@utdt.edu.

Previous studies suggest that entrepreneurial talent misallocation through these channels is crucial for explaining productivity differences between countries quantitatively. Indeed, it has been reported that correcting this mismatch can improve total factor productivity in less developed countries by as much as 25% (Hopenhayn, 2014).

The figure below shows the cumulative distributions of skills (literacy) for males aged 25 to 49 by employment status. In OECD countries, skill measures are quite ordered, in the sense that the distribution of skills among employees and self-employed dominates the same distribution for the unemployed and for those out of the labor force. On the other hand, the distribution of skills for the self-employed is no different from that for wage earners. The corresponding figures for Latin America contrast with the OECD ones. First, the distributions of skills for the unemployed and for employees are no different. Second, those out of the labor force have the highest distribution of skills. Third, the self-employed are much less skilled than employees, and even than the unemployed.

Figure 1: Skills by employment status



Notes: CDF of standardized literacy plausible value of males from 25 to 49 years of age by employment status and self-employment
 Source: PIAAC (OECD), STEP (Colombia and Bolivia) and ENHAB (Peru).

The table below shows the distribution of employed workers in OECD countries and in Latin America, according to whether they earn a wage or they are self-employed. About 40% of workers are self-employed, while this proportion is around 12% in selected OECD countries.

Table 1: Proportion of self-employed

	OECD	Latin America
Employees	87.6	59.4
Self-Employed	12.4	40.6

Notes: Males from 25 to 49 years of age.
 Source: PIAAC (OECD), STEP (Colombia and Bolivia) and ENHAB (Peru).

These figures suggest that the characteristics of the self-employed are very different in Latin America and in OECD countries. Low-skilled workers self-select more often into self-employment in Latin America. In other words, the allocation of human capital does not seem to favor the creation of highly productive firms. On the contrary, a high proportion of low-skilled workers seem to escape from unemployment through the creation of low-productivity, small and informal firms.

A few additional observations are in place. First, education quality is lower in Latin America, as compared to the OECD: any international assessment of skills for a given group (controlling for age, as in PISA data, or controlling for formal education using PIAAC and STEP data) shows that Latin American countries lag behind those in the OECD. Second, as we have already mentioned, the distributions of cognitive skills across occupations (employee or self-employed) and by employment status (employed, unemployed or out of the labor force) are less clearly ordered in the region than they are in OECD countries. As those skills are mainly acquired through schooling, these observations suggest that education may play a significant role when explaining entrepreneurial talent misallocation. In addition, the proportion of self-employed in Latin American countries is high, and the literature has shown that self-employment activity is not driven by innovation, but by low chances of finding a good job. In other words, many of those self-employed are replicative entrepreneurs in low productivity, small and informal firms.

One of the channels through which education influences worker allocation is provided by its signaling role in the labor market (Arozamena and Ruffo, 2015). Firms and workers are matched in the labor market on the basis of the information that education supplies on workers' abilities. If education's performance as a signaling device is worse -i.e if it provides coarser, less precise information to firms- the resulting allocation of skills to jobs yields lower productivity, fewer vacancies and a less efficient overall labor market performance.

In this paper, we analyze the signaling role of education and its impact on skills allocation in the labor market when workers can decide whether to work for a firm or to resort to self-employment. The possibility of self-employment changes the way education as a signal impacts labor market performance. If we concentrate exclusively on its signaling role, education is a means to inform potential employers about workers' skills. If the information provided is precise, then, as described in the previous paragraph, efficient talent and skill allocation in the labor market follows, which makes employment a relatively more attractive prospect. Then, only those with significant talents for self employment will skip the labor market. If the signal works poorly, however, the employment option becomes less luring, and a wider array of skill levels may resort to self-employment.

In a pure signaling setup, education does not affect workers' productivities, but only transmits information. In particular, it has no impact on a worker's ability to produce as self-employed. Needless to say, education is not only a signaling device, but also a way to acquire human capital. Then, demand for education depends on its future returns, as emphasized in human capital accumulation theory. Returns to education may be influenced by the availability of self-employment as an option. In particular, if the skills required to run a business are not provided by formal education, then workers will be less prone to invest in education.

In our framework, we combine these two roles for education, both as a way to accumulate human capital and as a signaling device. Both roles are connected for firms; what is relevant for relative labor demand is the expected skill level by education. On the other hand, as self-employed, the worker is only affected by the part of human capital accumulation that is useful in his independent job, but not by any information education may provide to others.

We then examine the possibility that education perform its roles with different degrees of effectiveness -i.e. the quality of education may be higher or lower. The impact of low educational

quality is twofold. First, it generates lower human capital accumulation. Second, workers with the same innate skills may choose different educational levels thus making education is less informative. Overall, the first effect may operate both on employment in the labor market and on self-employment, but the second only influences the prospects of working for a firm.

How will lower education quality impact labor market performance? How does it affect the pool of worker types that resort to self-employment? Our work will be mainly aimed at identifying the effect of low educational quality on skills allocation. In particular, we will analyze to what extent incorporating the self-employment option could change the effects of lower educational quality, as analyzed in the previous literature.

The next section describes the data on skills, education and self-employment that is used in the paper. Section 3 outlines the model, and provides a simple example to emphasize the impact of low educational quality on equilibrium behavior and labor market performance. Section 4 details the calibration procedure and Section 5 describes the main results that follow. Section 6 concludes. The Appendix includes further details on our data sources and on the calibration procedure.

2 Data on education, skills and self-employment

To examine the connections between education, skills and self-employment, we resort to standardized international assessment of skills, which provide a common methodology between countries. We use PIAAC results for OECD countries and STEP data for Latin American countries.² We also have information for Peru, from ENHAB. The methodology of this source, however, is not necessarily comparable with the measures derived from PIAAC and STEP. For that reason, we take literacy scores and standardize them by country (i.e. we subtract the mean and divide by the standard deviation). By making this transformation, typically used in the analysis of this type of variable, we avoid any difference in the mean and variance of literacy skills between PIAAC, STEP and ENHAB. It must be stressed, though, that the introduction of ENHAB does not change the results we want to emphasize.

We limit our analysis to males aged from 25 to 49. This selection is intended to abstract from influences that operate differently on various demographic groups. We also restrict the age range so that formal education decisions are already taken and retirement does not affect participation decisions. As before, our results are qualitatively robust to changes in these restrictions.

Figure 1 in the Introduction plots the c.d.f. of the standardized measures of literacy skills for selected OECD countries and Latin American countries. We group observations according to their employment status. In these plots, it is easy to see that differences in skills are on the whole distribution and not just on the mean.

Table 2 shows coefficients resulting from an OLS regression of literacy skills on dummy variables that characterize employment status, where employees is the base case. We also add controls for age, level of education, and number of years of education in columns (3) and (4). In columns (1) and (3) we show the results for OECD countries (from PIAAC). The coefficient of self-employment is not different from zero, implying that the self-employed's skills are not significantly different from those of employees. The coefficients of both unemployment and out of the labor force are negative and strong: skills in these groups are about half a standard deviation lower than those of employed

²PIAAC covers Austria, Belgium, Canada, Denmark, Estonia, Finland, France, Great Britain, Germany, Ireland, Italy, Japan, S.Korea, Netherlands, Norway, Poland, Russia, Slovakia, Spain, Sweden and the US. STEP countries are Bolivia and Colombia.

Table 2: Regressions of skills on employment status

	(1) OECD	(2) LA	(3) OECD	(4) LA
Self-Employed	0.02 (0.04)	-0.28*** (0.07)	0.02 (0.04)	-0.08 (0.06)
Unemployed	-0.52*** (0.06)	-0.13 (0.16)	-0.35*** (0.07)	-0.16 (0.15)
OLF	-0.57*** (0.06)	0.25 (0.16)	-0.32*** (0.07)	0.16 (0.14)
Constant	0.08*** (0.01)	0.28*** (0.05)	-0.75*** (0.07)	-0.71*** (0.09)
Controls	No	No	Yes	Yes
R^2	0.04	0.02	0.35	0.33
N	6,714	1,578	4,487	1,561

Notes: OLS estimations (standard errors) of standardized literacy scores on indicator variables of employment status, where the base case is employee. Additional controls include dummies by age, education and country. Sample is restricted to males from 25 to 49 years of age and with comparable levels of education between countries according to ISCED. *, **, and *** indicate significance at the 10%, 5% and 1% levels.

Source: PIAAC, STEP and ENHAB standardized literacy scores. Literacy scores are standardized by country.

workers. The unemployed's skills, just as those corresponding to workers that are out of the labor force, are lower even after including controls (see column 3), while differences between these two groups are smaller. In columns (2) and (4) we present analogous results Latin American countries (from STEP). Results are clearly different from those in OECD countries. The self-employed's skills are almost one third of a standard deviation lower than those of employees. Workers that are out of the labor force have higher skills than the self-employed, which is also striking. This difference is not significant, though. In addition, differences fall after including controls for age and education. This implies that most of the lower skill level of the self-employed in the Latin American countries is explained by lower education.

In addition, a main point of our model is related to the quality of education as a signaling device of skilled workers. In a related paper (Arozamena and Ruffo, 2015) we provide evidence suggesting that education does not perfectly signal worker's abilities, and that this problem is stronger in Latin America than in OECD countries.

The main indicator for education that we want to emphasize is the distribution of skills within each level of education. We then associate higher heterogeneity in abilities within each level of education with a less precise signal. Table 4 reports that the standard deviation of literacy scores within each level of education is higher in Latin America, even after controlling for country fixed effects.

A worker is not only productive because of cognitive skills, but also because of socio-emotional skills, as the literature has shown. If both types of skills are important, then, the quality of education should be measured by its performance in building and signaling both of them.

To examine this issue we use information from ENHAB, Peru. This source is rich in the diversity of cognitive and noncognitive measures it provides, and thus allows us to consider both types of skills. In order to provide simpler measures, we use exploratory factor analysis to build factors

Table 3: Probit of self-employment on measures of skills and controls

	(1) OECD	(2) LA	(3) OECD	(4) LA
Literacy	0.02 (0.02)	-0.18*** (0.05)	-0.11** (0.05)	-0.40*** (0.11)
Numeracy			0.16*** (0.05)	0.10 (0.10)
30 to 34 years	0.15* (0.09)	0.07 (0.14)	0.15* (0.09)	-0.26 (0.24)
35 to 39 years	0.35*** (0.08)	0.31** (0.14)	0.34*** (0.08)	-0.18 (0.26)
40 to 44 years	0.40*** (0.08)	0.39** (0.15)	0.39*** (0.08)	0.06 (0.26)
45 to 49 years	0.41*** (0.08)	0.43*** (0.16)	0.40*** (0.08)	0.59** (0.27)
Constant	-1.37*** (0.09)	-0.43*** (0.12)	-1.39*** (0.09)	-0.09 (0.17)
N	5,899	1,469	5,899	434

Notes: Probit estimations (standard errors) of the probability of being self-employed on standardized literacy and numeracy scores. Additional controls include dummies by age and country. Sample is restricted to males from 25 to 49 years of age. *, **, and *** indicate significance at the 10%, 5% and 1% levels.
Source: PIAAC, STEP and ENHAB standardized scores.

that summarize several variables. In particular, we estimate a single factor of cognitive skills that captures information from scores in literacy, numeracy, verbal fluency and working memory. Additionally we estimate a single factor of noncognitive skills, combining the scores of six measures of socioemotional abilities (Big Five and Grit).

In Arozamena and Ruffo (2015) we showed that the within variance by education is always higher for noncognitive skills compared to cognitive skills. In fact, noncognitive standard deviation within each level of education is at least 20% higher than the standard deviation of cognitive skills.

Finally, we found that education does not always imply an increase in skills in Peru. In particular, we found no difference in cognitive and noncognitive skills when we compare “some college” with “completed college.” Furthermore, the whole distribution of cognitive skills is indistinguishable between these two levels of education. This means that this change in the level of education does not imply a higher level of predicted cognitive skills. Importantly, this is not the case in OECD countries, where education and average skills are monotonically related for each level of education.

3 A model

To examine the issues described above we build a model combining two existing frameworks in the literature: (i) a labor-market setting with frictions, and (ii) a setup where education both enhances worker productivity through human capital accumulation and plays a signaling role. In addition, the model includes self-employment opportunities for workers, thus allowing us to examine the interplay between self-employment, labor market frictions and education decisions. In what follows,

Table 4: Standard deviation of literacy score

Education	(A)		(B)	
	OECD	Latin America	OECD	Latin America
Primary	49.10	77.72	48.42	70.63
High School	39.99	52.77	38.92	48.94
Tertiary	39.84	58.54	39.08	58.54
Some College	38.04	47.64	36.43	43.64
College	39.75	39.46	37.84	39.24
Total	48.47	70.80	40.07	56.61

Notes: Latin American countries are Colombia and Bolivia. Sample is restricted to individuals from 25 to 49 years of age and with comparable levels of education between countries according to ISCED.
Source: PIAAC and STEP literacy scores for selected countries.

we introduce each of these three elements in detail.

3.1 Basic setup

The structure of the basic model is as follows. There are n workers in the economy. Each worker has initial / innate skills denoted by $\theta \in [\theta_0, \theta_1]$. The value of θ is each worker’s private information, and cannot be directly observed by firms. Knowing his skills, each worker must choose an education level e . After having done so, he will learn which self-employment opportunity is available to him. Then, he must choose whether to take that opportunity or, alternatively, to enter the labor market expecting to receive wage offers. Both the value of self-employment opportunities and wage offers in the labor market will naturally depend on education. We describe educational choices, self-employment opportunities and the labor market’s operation next.

3.1.1 Education

Our approach is based on a signaling model, a very simple version of Spence (1973), extended to introduce idiosyncratic shocks to the cost of education. We allow education to increase a worker’s productivity through a human capital accumulation process.

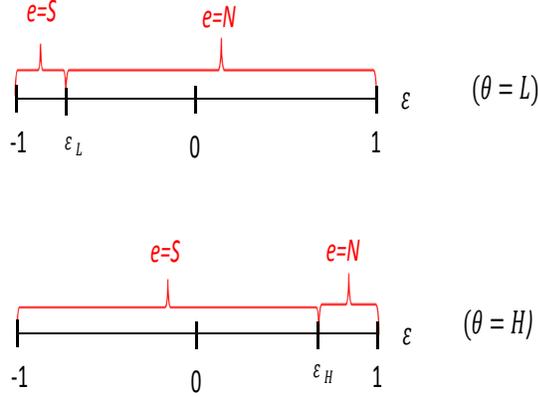
Prior to entering the labor market, each worker selects an education level $e \in \{N, S\}$ (no schooling, schooling). This choice will be the only worker characteristic that any firm will observe later on in the labor market, since his skill level θ is his private information. The worker’s utility is

$$u(e, \theta, \varepsilon) = Y(e, \theta) - c(e, \theta, \varepsilon),$$

where $Y(e, \theta)$ is the expected value of his future income, and the cost of education is

$$c(e, \theta, \varepsilon) = \begin{cases} \psi(\theta) + \sigma\varepsilon & \text{if } e = S \\ 0 & \text{if } e = N \end{cases},$$

where $\psi' < 0, \sigma > 0$ and ε is an i.i.d. random shock to costs. Then, the cost of education decreases with the worker’s ability.



The noise present in the cost function for education will be our instrument to evaluate how well education may play a signaling role in this economy. A worker with a high value of θ may choose no education -even if choosing a higher level of e is rewarded by the labor market- if ε is sufficiently high. A higher value of σ may then make education perform worse as a signaling device.

Assume, for instance, that $\varepsilon \sim U[-1, 1]$ and, for simplicity, that there are two skill levels, so that $\theta \in \{L, H\}$, with $H > L$. Furthermore, assume as well that $Y(e, \theta) = \hat{Y}(e)$ -i.e. unobservable worker skills are not rewarded in the labor market, but education is. Then, if

$$\psi(L) > \hat{Y}(S) - \hat{Y}(N) > \psi(H)$$

and σ is very small, there will be full separation of worker types according to education: all the highly skilled will educate, all those with low skills will not. As σ grows larger, though, such separation will eventually disappear. If $\varepsilon > \frac{\hat{Y}(S) - \hat{Y}(N) - \psi(H)}{\sigma} \equiv \varepsilon_H$, a highly-skilled worker will choose not to educate, while if $\varepsilon < \frac{\hat{Y}(S) - \hat{Y}(N) - \psi(L)}{\sigma} \equiv \varepsilon_L$, an unskilled worker would prefer education.

Then, at least for some values of σ , a larger value of σ generates less separation, thereby making education perform worse as a signal.

3.1.2 Self-employment

After making his educational choice, each worker learns which self-employment opportunity is available to him. Namely, the worker observes the realization of a random variable ϕ , distributed according to the c.d.f. $H(\phi)$, with support $[0, \bar{\phi}]$. His income from self-employment will be given by

$$f(e, \theta, \phi).$$

We assume that $f(\cdot)$ is increasing in all of its arguments (weakly increasing in e).

Once he observes ϕ the worker chooses whether to pursue his self-employment opportunity or, alternatively, to enter the labor market. Thus, the worker would decide to be self-employed whenever $f(e, \theta, \phi) > W(e)$, where $W(e)$ is the expected labor market income. His income from self-employment is perfectly known, whereas -as we will explain below- his labor-market income is stochastic, and he may even end up unemployed. We describe the functioning of the labor market next.

3.1.3 A labor market with frictions

Our approach to the labor market is based on a static version of Burdett and Mortensen (1998). We add heterogeneity to that version both among workers and among firms. This should allow us to examine the interplay between education -which may perform a signaling role- and the allocation of workers to vacancies in the labor market, both with and without self-employment opportunities.

There are two types of firms, W, B (say, white and blue collar). Let m_j be the number of type- j firms, $j = W, B$. The total number of firms is $m = m_W + m_B$.

As we have already mentioned, each worker that enters the labor market does so after having selected an education level e and foregoing his self-employment opportunity. Let n_e be the number of workers with education level e , $e = S, N$ ($n = n_S + n_N$), and $\hat{n}_e \leq n_e$ be the number of workers with education level e that have entered the labor market.

Firms and workers who have entered the labor market interact as follows. In order to employ a worker, a firm has to open a vacancy and post a wage offer associated with it. Hence, each firm must decide how many vacancies to open. Let v_{je} be the number of vacancies that a type- j firm creates for workers with education e , and let $v_j = v_{jN} + v_{jS}$. The cost of creating vacancies is given by $\kappa(v_j)$, with $\kappa', \kappa'' > 0$.

Let

$$p(\theta, j, e)$$

be the productivity of a worker of type θ with education e when employed by a type- j firm. We will thus assume that productivity varies across jobs, education and innate ability levels. In particular, productivity depends positively on innate skills and schooling. We will later set values so that high-skilled workers are more productive in type- W than in type- B firms.

Then, if a type- j firm posts an offer w_{je} for a worker with education e and that offer is taken, the filled job provides an expected profit

$$\pi_{je} = p_{je} - w_{je},$$

where $p_{je} = E[p(\theta, j, e)|e]$. The expectation is taken according to firms' beliefs about which types of worker select education level e and enter the labor market.

The labor market is thus segmented according to education. Workers of given education receive wage offers at random. Specifically, a wage offer intended for workers with education e will reach any given worker with probability $\frac{1}{\hat{n}_e}$. Then, the total number of offers that a given worker receives is distributed binomially. When the numbers of workers and firms are large, the binomial is approximated by a Poisson distribution with mean

$$\lambda_e = \frac{m_W v_{W_e} + m_B v_{B_e}}{\hat{n}_e}.$$

If a worker receives more than one offer, he chooses the highest wage. If he is reached by no offers, he stays unemployed and receives b as household production.

3.2 Equilibrium

We will now describe equilibrium behavior in our model. Let us start by assuming that \hat{n}_S (\hat{n}_N) educated (respectively, uneducated) workers have entered the labor market and examining firms' equilibrium behavior. Of course, firms anticipate the previous choices (education, self-employment) made by workers conditional on their private information, and update their beliefs about worker types accordingly.

3.2.1 Wage offers

Firms make their wage offers simultaneously and, as we mentioned above, those offers reach workers following a Poisson distribution with mean λ_e ($e = N, S$). This yields a mixed-strategy equilibrium that we now describe. As explained above, the labor market is segmented according to workers' educational levels. Let $F_e(w)$ be the equilibrium c.d.f. for wage offers in the segment corresponding to education e . In addition, let $P(F_e(w), \lambda_e)$ be the probability that a wage offer w be accepted in that segment (i.e. the probability that it be the largest offer received by a worker). We will examine equilibria that are symmetric among firms of the same type.

To describe the structure of the mixed-strategy equilibrium, we proceed by examining wage competition given vacancies, and then turn to firms' incentives to create vacancies anticipating the result of wage competition. Let us first assume that only one type of firm (say, type j) opens vacancies in the market for workers with education e . Therefore, $\lambda_e = \frac{m_j v_{je}}{\bar{n}_e}$. Given the number of vacancies, competition in wage offers among type- j firms cannot lead to a symmetric pure strategy equilibrium. Assume all firms offer the same wage w with certainty. Such an offer is not accepted with certainty, since the worker receiving it may be reached by more than one offer. Then, if $w < p_{je}$, any of the firms with vacancies in the market can deviate by offering $w + \delta$, a wage that will be accepted by a worker with probability one. For δ small enough, this deviation would be profitable. In addition, if $w = p_{je}$, firms make zero profits. Any firm may offer $w = b$ instead. Such an offer will be accepted by a worker that is reached by no other offers, which happens with positive probability -namely, $e^{-\lambda_e}$. The firm's expected profit will thus be positive, making the deviation profitable.³

In the mixed-strategy equilibrium, all offers in the support of $F_e(w)$ must yield the same expected profit. Let $[\underline{w}_{je}, \bar{w}_{je}]$ be that support.⁴ We cannot have $\underline{w}_{je} > b$, since \underline{w}_{je} is only accepted by a worker if he is reached by no other offers. In that case, the firm can lower the offer and still have it accepted when the worker has no alternatives. Thus, $\underline{w}_{je} = b$. This pins down equilibrium expected profits for firms, since offering $\underline{w}_{je} = b$ yields $(p_{je} - b)e^{-\lambda_e}$. All other wages in the support must generate the same profits, which implies that \bar{w}_{je} , which is accepted with certainty, is such that $p_{je} - \bar{w}_{je} = (p_{je} - b)e^{-\lambda_e} > 0$. Then,

$$\bar{w}_{je} = p_{je}(1 - e^{-\lambda_e}) + be^{-\lambda_e}$$

As regards the mixed strategy that firms use in the symmetric equilibrium, note that, given the process by which offers reach workers, the probability that an offer w be accepted is

$$P(F_e(w), \lambda_e) = \sum_{x=0}^{\infty} F_e(w)^x \frac{e^{-\lambda_e} \lambda_e^x}{x!} = e^{-\lambda_e [1 - F_e(w)]}.$$

Given that all offers must yield the same expected profit, then,

$$(p_{je} - w)e^{-\lambda_e [1 - F_e(w)]} = (p_{je} - b)e^{-\lambda_e}$$

for all $w \in [b, \bar{w}_{je}]$. Thus,

$$F_e(w) = \frac{1}{\lambda_e} \ln \left(\frac{p_{je} - b}{p_{je} - w} \right).$$

³An analogous argument rules out the possibility that some employers make the same wage offer with certainty. A fuller description of the equilibrium may be found in Mortensen (1990).

⁴It can be shown that $F_e(w)$ must be continuous and have a connected support.

Still taking as given the number of vacancies, assume now firms of both types, W and B , created vacancies for workers with education e , so that $\lambda_e = \frac{m_W v_{W_e} + m_B v_{B_e}}{\hat{n}_e}$ with $v_{W_e}, v_{B_e} > 0$. In line with what will happen in the numerical exercises later on, assume that $p_{W_e} > p_{B_e}$ for $e = S$. By an argument analogous to the one used above, any equilibrium that is symmetric within each firm type must be in mixed strategies.

Take $F_e(w)$ as the distribution of wages in the market for workers with education e . Any wage w_{je} that is offered by a type- j firm must maximize its expected profits, $(p_{je} - w_{je})e^{-\lambda_e[1-F_e(w_{je})]}$. Then, taking any wages in the support of each type of firm's wage offer distribution, w_{W_e}, w_{B_e} , we must have

$$\begin{aligned} (p_{W_e} - w_{W_e})e^{-\lambda_e[1-F_e(w_{W_e})]} &\geq (p_{W_e} - w_{B_e})e^{-\lambda_e[1-F_e(w_{B_e})]} \\ &> (p_{B_e} - w_{B_e})e^{-\lambda_e[1-F_e(w_{B_e})]} \\ &\geq (p_{B_e} - w_{W_e})e^{-\lambda_e[1-F_e(w_{W_e})]} \end{aligned}$$

Combining these inequalities, it follows that we must have $w_{W_e} > w_{B_e}$ if those wages are in the support of each type of firm's equilibrium wage distribution. Therefore, those supports must be disjoint -except for the possibility that the maximum wage offer in type B 's support coincides with the minimum wage in type B 's.

Let $[\underline{w}_{B_e}, \bar{w}_{B_e}]$ ($[\underline{w}_{W_e}, \bar{w}_{W_e}]$) be the support of the wage offer distribution for type- B (respectively, type W) firms. Replicating our argument above for the case where there is only one firm type with open vacancies, we must have $\underline{w}_{B_e} = b$. In addition, $\bar{w}_{B_e} = \underline{w}_{W_e}$, since otherwise a type- W firm would increase its profits by lowering a wage offer \underline{w}_{W_e} to a slightly lower offer $\underline{w}_{W_e} - \delta$, which would be accepted with the same probability.

Replicating our previous argument for the case of one active firm type, we must have

$$F_e(w) = \begin{cases} \frac{1}{\lambda_e} \ln \left(\frac{p_{B_e} - b}{p_{B_e} - w} \right) & \text{if } w \in [b, \bar{w}_{B_e}] \\ \frac{1}{\lambda_e} \ln \left(\frac{p_{W_e} - \bar{w}_{B_e}}{p_{W_e} - w} \right) & \text{if } w \in [\underline{w}_{W_e}, \bar{w}_{W_e}] \end{cases}$$

where $\bar{w}_{B_e} = \underline{w}_{W_e}$. Given that the fraction of wage offers made by type- B firms is $\frac{m_B v_{B_e}}{m_B v_{B_e} + m_W v_{W_e}}$, we must have

$$\begin{aligned} F_e(\bar{w}_{B_e}) &= \frac{m_B v_{B_e}}{m_B v_{B_e} + m_W v_{W_e}} \\ F_e(\bar{w}_{W_e}) &= 1. \end{aligned}$$

As all wage offers in any firm type's support must yield the same expected profits, we have

$$p_{W_e} - \bar{w}_{W_e} = (p_{W_e} - \underline{w}_{W_e})e^{-\lambda_{W_e}}$$

and

$$(p_{B_e} - \bar{w}_{B_e})e^{-\lambda_{B_e}} = (p_{B_e} - b)e^{-\lambda_e},$$

where $\lambda_{je} = \frac{m_j v_{je}}{\hat{n}_e}$

3.2.2 Wage offers and vacancies

Let us now turn to the simultaneous determination of vacancies and wage offers. A type- j firm must select how many vacancies to create for each education level and how to make the wage offers

associated to each of them. That is, it has to solve

$$\max_{\substack{w_{jS}, w_{jN} \geq b \\ v_{jS}, v_{jN} \geq 0}} \pi_{jS} P(F_S(w_{jS}), \lambda_S) v_{jS} + \pi_{jN} P(F_N(w_{jN}), \lambda_N) v_{jN} - \kappa(v_{jS} + v_{jN}).$$

As all firms of the same type have the same productivity in each market and, as described above, wage offer equilibria will be symmetric among firms of any given type, firms would set wages in a way that generates the same expected profit for each type in the case of each education level. Let the equilibrium expected profit in for a type- j firm hiring a worker with education e be $\pi_{je}^*(p_{je})$. Thus, the firm's problem becomes

$$\max_{v \geq 0} \{ \pi_{jS}^*(p_{jS}) v_{jS} + \pi_{jN}^*(p_{jN}) v_{jN} - \kappa(v_{jS} + v_{jN}) \}$$

so than

$$\kappa'(v_{jS} + v_{jN}) = \max \{ \pi_{jS}^*(p_{jS}), \pi_{jN}^*(p_{jN}) \} \quad (1)$$

This condition implies that the equilibrium profit from opening a vacancy in each market (for educated or uneducated workers) should be equal if both $v_{je} > 0$. This equality holds only for interior solutions, and would not hold if some $v_{je} = 0$. In other words, a firm can find it convenient to open vacancies just for one type of market. On the other hand, there could be equilibria where two productivities coexist for a given education level (according to the different firm types).

3.2.3 Employment and mean wages

A worker that participates in this labor market can be finally employed at different wages, or even unemployed. If unemployed, the worker receives the income of household production b ; if employed he expects to receive a mean wage that we need to compute. For doing so it is important to consider that the distribution of paid wages is different from the distribution of wage offers. The reason is that wages are not accepted at random by workers, but, on the contrary, low wages are more frequently rejected. Thus, the distribution of accepted and paid wages is the product of the distribution of wage offers times the probability of acceptance of these wage offers, that can be defined as:

$$G_e(w) = \begin{cases} \frac{1}{\lambda_e} \ln \left(\frac{p_{Be} - b}{p_{Be} - w} \right) e^{[\ln(p_{Be} - b) - \ln(p_{Be} - w) - \lambda_e]} & \text{if } w \in [b, \bar{w}_{Be}] \\ \frac{1}{\lambda_e} \ln \left(\frac{p_{We} - \bar{w}_{Be}}{p_{We} - w} \right) e^{[\ln(p_{We} - \bar{w}_{Be}) - \ln(p_{We} - w) - \lambda_e]} & \text{if } w \in [\underline{w}_{We}, \bar{w}_{We}] \end{cases}.$$

From this distribution we can compute the mean wages of workers if they are finally employed at the high type and low type firm as

$$\begin{aligned} \omega_{We} &= p_{We} - (p_{We} - \bar{w}_{Be}) e^{-\lambda_{We}} \left(1 + \frac{\lambda_{We}}{2} \right) \\ \omega_{Be} &= p_{Be} - (p_{Be} - b) e^{-\lambda_{Be}}, \end{aligned}$$

respectively. Finally, notice that the proportion of workers with education e that are unemployed is $e^{-\lambda_e}$, those employed in a high-type firm are those that receive at least one offer from a firm of that type, $(1 - e^{-\lambda_{We}})$, and those employed in a low-type firm are those that do not receive a wage offer from a high-type firm but receive at least one from a low-type, firm, $e^{-\lambda_{We}} (1 - e^{-\lambda_{Be}})$. Finally, the expected income of a worker with education e is

$$W(e) = (1 - e^{-\lambda_{We}}) \omega_{We} + e^{-\lambda_{We}} (1 - e^{-\lambda_{Be}}) \omega_{Be} + b e^{-\lambda_e}.$$

3.2.4 Self-employment

As described above, and given his education e , once he observes ϕ the worker must decide whether to enter the labor market or turn to self-employment. As self-employed he would receive $f(e, \theta, \phi)$, while in the labor market he is subject to the risk of not finding any job. Given that we consider risk-neutral agents, workers will decide to be self-employed whenever $f(e, \theta, \phi) > W(e)$, that is whenever the income as self-employed is higher than the expected income in the labor market for a worker with education e . We describe this with the following decision rule.

$$Q(e, \theta, \phi) = \begin{cases} 1 & \text{if } f(e, \theta, \phi) > W(e) \\ 0 & \text{if } f(e, \theta, \phi) < W(e) \end{cases} \quad (2)$$

3.2.5 Education

Thus, the expected labor income of choosing education e is

$$Y(e, \theta) = E \max \{f(e, \theta, \phi), W(e)\}$$

where the expectation is taken with respect to ϕ . This expression follows from the timing of the model, where worker first selects his level of education, then he observes ϕ and decides later whether to be self-employed. All of this happens before his participates in the labor market, if he chooses to do so.

Clearly, then, and given θ , a worker will select an education level e^* such that

$$e^* = \arg \max_e \{Y(e, \theta) - c(e, \theta, \varepsilon)\}. \quad (3)$$

More explicitly, we have,

$$e(\theta, \varepsilon) = \begin{cases} S & \text{if } Y(S, \theta) - Y(N, \theta) > \psi(\theta) + \sigma\varepsilon \\ N & \text{if } Y(S, \theta) - Y(N, \theta) < \psi(\theta) + \sigma\varepsilon \end{cases} \quad (4)$$

The noise present in the cost function for education will be our instrument to evaluate how well education may play a signaling role in this economy. A worker with a high value of θ may choose a low level of education -even if choosing a higher level of e is rewarded by the labor market- if ε is sufficiently high. A larger value of σ may then make education perform worse as a signaling device.

An equilibrium will then be given by an education decision rule, $e(\theta, \varepsilon)$, a self-employment decision rule $Q(e, \theta, \phi)$, wage offer distributions for both types of firms and vacancy choices, and beliefs by firms in the labor market such that (i) $e(\theta, \varepsilon)$ satisfies (4), (ii) $Q(e, \theta, \phi)$ satisfies (2), (iii) wage offer distributions and vacancy choices form a labor market equilibrium, and (iv) beliefs by firms are updated according to observed worker behavior, so that

$$p_{je} = E[p(\theta, j, e) | e, Q = 0]$$

3.3 An example

As an example of how the model works and so as to examine its equilibrium, we present now an example that includes a few additional assumptions.

Specifically, we assume that $p_{WN} \leq p_{BN} \leq p_{BS} \leq p_{WS}$. In a frictionless labor market, then, uneducated (educated) workers should be employed by type- B (respectively, type- W) firms. In the (hypothetical) case where productivities are such that both firm types open vacancies for both education levels, equilibrium wage and vacancy choices could be described as follows.

For each labor market, as defined by each education level, high productivity firms would have higher profits and offer higher wages. Thus, we have

$$\begin{aligned}\pi_B^*(p_{BN}, p_{BS}) &= \max \{e^{-\lambda_S} (p_{BS} - b), e^{-\lambda_{BN}} (p_{BN} - \bar{w}_{WN})\} \\ \pi_W^*(p_{WN}, p_{WS}) &= \max \{e^{-\lambda_N} (p_{WN} - b), e^{-\lambda_{WS}} (p_{WS} - \bar{w}_{BS})\}\end{aligned}$$

which means that firms of type W offer wages beginning at b for workers without education and firms of type B offer wages beginning at b for workers with education.

Then, for worker type S ,

$$\begin{aligned}F_S(\bar{w}_{BS}) &= \frac{1}{\lambda_S} \ln \left(\frac{p_{BS} - b}{p_{BS} - \bar{w}_{BS}} \right) = \frac{m_B v_{BS}^*}{m_B v_{BS}^* + m_W v_{WS}^*} \\ F_S(\bar{w}_{WS}) &= \frac{1}{\lambda_S} \left[\lambda_{SB} \ln \left(\frac{p_{WS} - \bar{w}_{BS}}{p_{WS} - \bar{w}_{WS}} \right) \right] = 1,\end{aligned}\tag{5}$$

while for worker type N

$$\begin{aligned}F_N(\bar{w}_{WN}) &= \frac{1}{\lambda_N} \ln \left(\frac{p_{WN} - b}{p_{WN} - \bar{w}_{WN}} \right) = \frac{m_W v_{WN}^*}{m_W v_{WN}^* + m_B v_{BN}^*} \\ F_N(\bar{w}_{BN}) &= \frac{1}{\lambda_N} \left[\ln \left(\frac{p_{BN} - \bar{w}_{WN}}{p_{BN} - \bar{w}_{BN}} \right) \right] = 1,\end{aligned}\tag{6}$$

where

$$\lambda_e = \frac{m_B v_{Be}^* + m_W v_{We}^*}{n_e}$$

Using these equations, the maximum wages \bar{w} can be expressed as

$$\begin{aligned}\bar{w}_{BS} &= p_{BS} (1 - e^{-\lambda_{SB}}) + b e^{-\lambda_{SB}} \\ \bar{w}_{WS} &= p_{WS} (1 - e^{-\lambda_{SW}}) + \bar{w}_{BS} e^{-\lambda_{SW}} \\ \bar{w}_{WN} &= p_{WN} (1 - e^{-\lambda_{NW}}) + b e^{-\lambda_{NW}} \\ \bar{w}_{BN} &= p_{BN} (1 - e^{-\lambda_{NB}}) + \bar{w}_{WN} e^{-\lambda_{NB}}\end{aligned}$$

where

$$\lambda_{je} = \frac{m_j v_{je}^*}{\hat{n}_e}$$

Then, combining the FOCs in (1) and the wage distributions we have four equations implicitly determining the endogenous variables v_{je}^* .

To simplify expressions, we will assume that educated workers would be hired only by type- W firms, while uneducated workers would be employed only by type- B firms.⁵ In such a case, the

⁵It can be shown that, in this model, if one type of firm opens vacancies for both types of workers, the other type of firm will open vacancies just for one type of worker or will not open vacancies at all.

expected wage is given by

$$\begin{aligned}\omega_S &= p_{WS} - (p_{WS} - b) e^{-\lambda_S} \left(1 + \frac{\lambda_S}{2}\right) \\ \omega_N &= p_{BN} - (p_{BN} - b) e^{-\lambda_N} \left(1 + \frac{\lambda_N}{2}\right),\end{aligned}$$

and expected income is

$$\begin{aligned}W(S) &= b + (p_{WS} - b) (1 - e^{-\lambda_S}) \left[1 - e^{-\lambda_S} \left(1 + \frac{\lambda_S}{2}\right)\right] \\ W(N) &= b + (p_{BN} - b) (1 - e^{-\lambda_N}) \left[1 - e^{-\lambda_N} \left(1 + \frac{\lambda_N}{2}\right)\right].\end{aligned}$$

This expected value is important to understand workers' decisions when choosing between entering the labor market and self-employment.

We will assume that a worker's productivity as self-employed has the following properties: $\frac{\partial f(e, \theta, \phi)}{\partial e} = 0$, $\frac{\partial f(e, \theta, \phi)}{\partial \theta} = 0$, $\frac{\partial f(e, \theta, \phi)}{\partial \phi} = 1$ and that the distribution of $\phi \in [0, \bar{\phi}]$ is uniform. We define $\hat{\phi}_{e\theta}$ as the reservation productivity value above which workers with education e and skills θ decide to be self-employed. This value is the implicit solution to

$$f(e, \theta, \hat{\phi}_{e\theta}) = W(e).$$

This implies that the proportion of workers with education e and skills θ that will decide to enter the labor market will be $H(\hat{\phi}_{e\theta})$. Given our assumptions on $f(\cdot)$ in this example, we should have $f(e, \theta, \phi) = \phi + f_0$, where f_0 is a constant. For simplicity, let us assume that $f_0 = 0$.

Let us rewrite the expected income of a worker with education e and skill level θ before the shock ϕ is realized as

$$\begin{aligned}Y(e, \theta) &= E[\max\{f(e, \theta, \phi), W(e)\}] \\ &= H(\hat{\phi}_{e\theta})W(e) + [1 - H(\hat{\phi}_{e\theta})] E[f(e, \theta, \phi) | \phi > \hat{\phi}_{e\theta}] \\ &= \frac{\hat{\phi}_{e\theta}}{\bar{\phi}} W(e) + \frac{\bar{\phi}^2 - \hat{\phi}_{e\theta}^2}{2\bar{\phi}}\end{aligned}$$

using the fact that $H(\cdot)$ is uniform on $[0, \bar{\phi}]$. Accordingly, a type- θ worker will choose schooling whenever

$$Y(S, \theta) - c(S, \theta, \varepsilon_i) > Y(N, \theta),$$

We can rewrite the education condition as

$$c(S, \theta, \varepsilon_i) < \frac{\hat{\phi}_{S\theta}}{\bar{\phi}} W(S) + \frac{\bar{\phi}^2 - \hat{\phi}_{S\theta}^2}{2\bar{\phi}} - \left[\frac{\hat{\phi}_{N\theta}}{\bar{\phi}} W(N) + \frac{\bar{\phi}^2 - \hat{\phi}_{N\theta}^2}{2\bar{\phi}} \right] \quad (7)$$

$$= [W(S) - W(N)] \frac{\hat{\phi}_{N\theta}}{\bar{\phi}} + \left[W(S) - \frac{\hat{\phi}_{S\theta} + \hat{\phi}_{N\theta}}{2} \right] \frac{\hat{\phi}_{S\theta} - \hat{\phi}_{N\theta}}{\bar{\phi}} \quad (8)$$

The inequality in (7) and (8) is intuitive since it states that what matters for investing in education is the increase in labor market income that it brings multiplied by the probability of participating in the labor market. More specifically, since, for any θ , we must have $\hat{\phi}_{S\theta} > \hat{\phi}_{N\theta}$, the first term in (8) is given by the probability that the worker is employed in the labor market both with and without schooling (namely, $\hat{\phi}_{N\theta}/\bar{\phi}$) multiplied by the expected labor income rise that education generates for the worker. The second term is given by the probability that the worker is employed if educated but resorts to self-employment if uneducated (i.e. $(\hat{\phi}_{S\theta} - \hat{\phi}_{N\theta})/\bar{\phi}$) multiplied by the expected income increase education gives rise to -the worker would receive ϕ as self-employed if $e = N$, but gets expected labor income $W(N)$ if educated. With probability $(\bar{\phi} - \hat{\phi}_{S\theta})/\bar{\phi}$, though, the worker is self-employed both with and without education, so schooling generates no additional expected income. Education, then, is more valuable when the probability that the worker participates in the labor market is larger.

In other words, the higher the chances that a worker has to be self-employed the lower the incentives to invest in education. Of course, this result is driven by the extreme assumptions of this section, in particular, that education and skills do not have any impact on the productivity as self-employed. Nevertheless, In any case, education has a signaling role only in the labor market, but not for self-employment. For that reason, whenever the signaling motive is relevant, the higher are the ex-ante chances of being self-employed, the lower the incentives to educate will be.

4 Calibration

This section describes the calibration of the model, choosing parameters in line with Peruvian data. We first describe our parametric choice and discuss its identification. Then, we present the data moments and the fit of our calibrated model.

4.1 Parametrization

We set skills to two values, θ_L and θ_H . We consider that education changes these innate skills by increasing the worker's production by a factor of $(1 + \gamma)$. We set job productivities as follows:

$$\begin{aligned} p_{jS} &= p_{jH}\theta_H(1 + \gamma)x_{HS} + p_{jL}\theta_L(1 + \gamma)(1 - x_{HS}) \\ p_{jN} &= p_{jH}\theta_H x_{HN} + p_{jL}\theta_L(1 - x_{HN}) \end{aligned}$$

where x_{He} is the proportion of workers with skills θ_H with education e in the labor market and where p_{jH}, p_{jL} are productivities of a given type of worker in a given firm.

We parametrize our cost function for vacancies as $\kappa(v) = av^\alpha$, where a is a level parameter and α is an elasticity parameter. The productivity function of a self-employed worker is set to $f(\theta, \phi) = f_0 + f_1\theta + \phi$, where $\phi \sim U[0, \bar{\phi}]$.

Finally, we the cost function for education is $c(S, \theta, \varepsilon) = c_0 - c_1\theta + \sigma\varepsilon$, with $\varepsilon \sim U[-1, 1]$.

4.2 Identification

We now turn to parameter identification in this version of the model. Here, we provide intuitive arguments for parameter identification and we refer the reader to the Appendix for formal details. We present the parameters to calibrate in groups, and then relate them to particular outcomes.

We exploit moments related to the education decision, such as the proportion of educated workers, and the proportion of educated workers by skills, to identify parameters c_0 , c_1 and σ . These parameters are related to the cost of education. Fixing income, an increase in c_0 reduces education for both skills; increasing c_1 induces education of the high skilled workers more; a higher value of σ implies that costs of high and low skilled workers tend to overlap more. Thus, these parameters are related to the level and separation in education.

To identify parameters f_0 , f_1 and $\bar{\phi}$, we consider the proportion of entrepreneurs by education and the skill productivity differential among the self-employed. In fact, a higher value of f_0 implies higher incentives to become self-employed for all workers; f_1 increases these incentives only for highly skilled workers, and increases the skill productivity differential; a higher value of $\bar{\phi}$ raises productivity for the self-employed without changing the productivity differential.

As for the identification of employees productivity values, we use both income differentials and the proportion of workers in each job market. In particular, we consider the following correlation between worker characteristics and wages: the skill wage differential, the education wage differential and the white collar wage differential. These would identify θ_H , γ and p_{je} . Productivity values will also determine the proportion of educated workers in type- W and in type- B jobs, and the proportion of non-educated workers in each type of job.

Finally, we use a very simple version of a labor market model, with only one type of worker and without the option to self-employ, to show that the unemployment rate and the elasticity of that rate to productivity identify the crucial parameters of the cost function (see Appendix). In particular, given productivity and household production, b , the unemployment rate identifies the level parameter of the vacancy cost function, a , while the elasticity of unemployment with respect to productivity identifies the elasticity parameter, α . Then, we use the unemployment rate by education, and the unemployment elasticity with respect to productivity as targets.

4.3 Moments for calibration: data

The model is calibrated in line with the Peruvian data from ENHAB for workers aged between 25 and 40. We focus on this age group since young workers are the most affected if education is an inaccurate signal. In addition, we drop all inactive individuals, i.e. those without attachment to the labor market.

We intend to use the model to represent the case where education's inaccuracy as a signal yield partial pooling in equilibrium. Then, we attempt to represent the correlation of education and skills as in the data.

In the model, we distinguish between two education levels.

The ENHAB survey is particularly interesting for our purposes because it explores different relevant worker skill dimensions, such as literacy, numeracy, and socio-emotional skills. Using a principal components approach we condense this information to a single indicator of skills, and we set a number in this distribution to divide the population in two groups: high- and low-skilled. Additionally, we classify workers as self-employed or employees, and jobs as blue- or white-collar exploiting the classification of occupations. Finally, we identify workers according to their education by considering educated all workers with completed college or tertiary education, or higher.

Using these classifications, we compute the composition of workers. For example, we compute the proportion of skilled workers among those educated and the proportion of skilled workers among those without tertiary or college degrees. These figures are presented in Table 6. They show that workers with higher skills are also more educated, but that the proportion of uneducated, highly

skilled workers is also large. Self-employment is higher for those with no education. In particular, the proportion of self-employment is around 27% for educated workers while it is 49% among workers without high education. The remaining workers are mainly in white-collar jobs, which comprise about 53% among wage earners. Unemployment is about 7% for workers with high education while it is almost 11% for uneducated workers.

In addition to analyzing worker proportions, we also consider education, skills and occupation wage differentials. We consider first wage earners, and run a regression of the log of wages on our binary indicator of education. We find that high education is related to 50% higher wages. Additionally, the analogous regression on skills and white-collar occupation yield a wage differential of 32% and 36%, respectively. By restricting the sample to those self-employed, we examine the education labor income differential. We run a regression of the log of labor income on our binary variable of education and find a differential of 39%. We interpret this as an effect of skills rather than of education. In fact, when we run the same regression including skills and education, we find that the coefficient for education lacks any significance, while the coefficient for skills implies a labor income differential of 54%.

Finally, we exploit the time series from the Instituto Nacional de Estadísticas e Informática (INEI) on employees, GDP, productivity, and unemployed, from 2001 to 2011. During this period, productivity rose about 47%, while unemployment fell by 26%. We use this information to compute the elasticity of unemployment with respect to productivity, which takes a value of -0.55.⁶

4.4 Calibrated parameters

Table 5 presents all parameter values, both those that we initially set and those that are chosen to target data moments.

First, we set $\theta_L = 1$ as a normalization. All other productivity parameters are the proportional to that normalized value. We also normalize the total number of workers to 2, and we set the distribution of skills as defined by our data so that $n_L/n_H = .75$.

A relevant parameter in the model is the number of firms per worker. Once the number of workers is set, the number of firms should be calibrated. We take the stand to calibrate this number to generate firms size that represent the median worker in the survey. The median employed worker is in a firm that hires 6 to 10 workers. Additionally, we observe that white-collar jobs are in larger firms than blue-collar jobs -specifically, the former firms being about twice the size of the latter. Thus, we set $m_W = 1/5, m_B = 1/10$.⁷

The remaining parameters that were calibrated to reproduce data moments as presented in Table 6. The calibrated model matches well the distribution of workers in education and in self-employment, as well as unemployment levels. Nevertheless, it fails to match some labor income differentials. In particular, the skills wage differential and the white-collar wage differential do not fit the ones observed in the data.

4.5 Algorithm

With productivity and vacancy cost function parameters and an initial value for the number of workers in each education level, \hat{n} , we can solve for the labor market problem and, among other

⁶We provide more details about the data in the Appendix.

⁷These numbers are not related to the firm size distribution in other surveys, suggesting that alternative calibrations of m could be even lower. This would be inconsequential, after recalibrating the cost function. See the Appendix for the implications of changes in m and, thus, in optimal v .

Table 5: Parameter values

Parameter	Description	Value
θ_L	Low skills, normalized to 1	1
θ_H	High skills	1.9
γ	Effects of education on labor market productivity	0.48
p_{BN}	Productivity parameter in firm type $j = B$ and education $e = N$	3.59
p_{BS}	Productivity parameter in firm type $j = B$ and education $e = S$	2.81
p_{WN}	Productivity parameter in firm type $j = W$ and education $e = N$	2.81
p_{WS}	Productivity parameter in firm type $j = W$ and education $e = S$	5.96
b	Household production	1.14
a	Vacancy cost function, level parameter	0.0001
α	Vacancy cost function, elasticity parameter	4.3
f_0	Self-employment productivity, level parameter	-6.05
f_1	Self-employment productivity, skills parameter	3.55
$\bar{\phi}$	Self-employment productivity, idiosyncratic productivity	5.69
c_0	Education cost function, level parameter	5.71
c_1	Education cost function, skills parameter	2.05
σ	Education cost function, dispersion of idiosyncratic cost	2.71
m_W	Number of white collar firms	0.1
m_B	Number of blue collar firms	0.2

outcomes, the expected labor market income for each education level, $W(e)$. Using these values and the self-employment productivity function $f(\theta, \phi)$, we can compute the labor-supply decision, the self-employment rate and expected labor income by type and education, $Y(\theta, e)$. Using the education cost function, $c(e, \theta, \varepsilon)$, we solve for the education decision, i.e. the thresholds in ε . This description suggests an algorithm for the quantitative solution of the model.

Let χ be a vector of eight positions describing the proportion of workers of each type, H and L , with the proportion of educated workers within each type and with the proportion of workers in self-employment within each of the four groups. To solve the model numerically we can define a mapping from χ into itself with the following algorithm:

1. Guess a value of χ^i , say one in which there is separation in education and no self-employment.
2. Given χ , solve for the labor market endogenous variables v^* and compute $W(e)$.
3. Given $W(e)$, solve for $\hat{\phi}_{e\theta}$ and $Y(\theta, e)$
4. Given $Y(e, \theta)$, solve for $e^*(\theta, \varepsilon)$
5. Given the solutions $e^*(\theta, \varepsilon)$ and $\hat{\phi}_{e\theta}$ compute $\hat{\chi}^i$.
6. For a fixed relaxation parameter $g \in (0, 1)$, compute a new guess of χ from $\chi^{i+1} = g\hat{\chi}^i + (1 - g)\hat{\chi}^i$.
7. Iterate on this scheme to convergence.

Table 6: Quantitative model: comparison with the data

Moment	Model	Data
Proportion of educated workers	0.2993	0.29
Proportion of educated workers with high skills	0.7273	0.73
Proportion of uneducated workers with high skills	0.5049	0.48
Proportion of educated workers that are self-employed	0.1818	0.23
Proportion of uneducated workers that are self-employed	0.4096	0.37
Education wage differential	1.6941	1.51
White collar wage differential	1.6941	1.36
Skills wage differential	1.1901	1.32
Education labor income differential (self-employed)	1.3038	1.39
Skills labor income differential (self-employed)	1.4355	1.54
Unemployment of educated workers	0.078	0.07
Unemployment of uneducated workers	0.088	0.11
Proportion of white collar jobs	0.53	0.532
Proportion of white collar jobs with educated workers	0.32	0.404
Unemployment elasticity wrt productivity	-0.55	-0.55

The algorithm proved to be robust, in the sense that the vector χ converges relatively fast and smoothly.

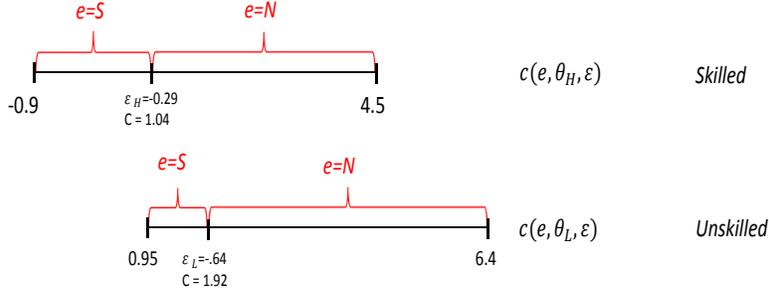
5 Results

We now turn to the results of the calibrated model. In the first column of table 7 we present the results of our baseline case. The calibrated parameters imply a solution in which there is partial pooling in education, meaning that not all workers with skills H are educated and some workers with skills L educate. The proportion of educated workers is 0.3. These results are reported in the first three lines. This distribution of workers in education imply partial pooling for both education groups.

Figure 5 shows education costs for both worker types. Education is less costly for the median skilled than for the median unskilled, but there is a strong overlap. This overlap is related to the idiosyncratic component of the costs of education, σ . The plot includes as well workers' decisions in this calibrated economy. It shows that the income effect of education is higher for the skilled than for the unskilled. This reinforces cost differences, and provides the skilled with higher incentives to educate.

The self-employment rate is 34% in our baseline economy. Significantly, the self-employment rate among educated workers is 18% while for the uneducated workers is 41%. This high proportion of self-employment reduce incentives to educate, as explicitly shown in equations (7) and (8). Particularly, the self-employment rate is higher among the unskilled, which implies a lower income effect of education for those workers.

It is important to emphasize that the allocation in self-employment is at the cost of lower productivity. In particular, for the unskilled and uneducated, the highest productivity in self-



employment is lower than productivity when working for a firm. The reason why unskilled and uneducated workers are self-employed is the labor market friction that generates lower wages and unemployment.

On average, employees have a productivity of 4.22, while this productivity is 3.49 for uneducated workers and 5.76 for educated ones. The average productivity among the self-employed is 4.55 (an average of 5.66 for educated workers and 4.34 for uneducated workers). In computing these averages we exclude the unemployed. A different computation is carried out for per capita productivity, where the unemployed are considered household workers, generating b units of product at home. Average per capita productivity is 4.04 (4.68 for educated workers and 3.76 for uneducated workers).

Workers are paid wages lower than productivity. Mean wages paid are $\omega_S = 4.9$ (educated) and $\omega_N = 2.9$ (uneducated). When including those self-employed, mean labor income by education is 4.62 and 2.75. The unemployment rate is 8.7% for all workers, 7.8% for those educated and 8.8% for those uneducated.

Column (2) in table 7 presents the results of the same quantitative model, but this time excluding self-employment: it follows from setting $f(\theta, \phi) = 0$. We do not consider this as a relevant counterfactual. We focus on this case to gain intuition of the main economic mechanisms behind our model. Without the choice of entrepreneurship, workers find more incentives to educate. Avoiding education is less tempting now, given that in possible employment paths there are positive return to education in equilibrium. Equations (7) and (8) show this relationship between the self-employment rate and incentives to educate.

Without self-employment, the labor market is more populated with workers, and firms get their vacancies filled with higher probability, thus increasing incentives to create vacancies. This raise in vacancies does not cover all new workers. On the contrary, unemployment increases. This is the result of frictions in the labor market. We show in the Appendix that the calibrated cost function implies an elasticity of vacancy creation with respect to the number of workers in a market that is smaller than one. The unemployment rate, that workers cannot avoid by self-employment, leads to a wider gap between productivity and average wages, which are now 4.62 and 2.17 for educated and uneducated workers, respectively. Overall, per capita productivity is 3.41, lower than the previous case.

The high proportion of uneducated and unskilled self-employed makes Latin America distinct from OECD countries analyzed, and the high share of self-employment among uneducated workers is probably related to the inefficiency or to the low quality of education. We analyze this issue by changing human capital accumulation and the dispersion in idiosyncratic costs of education, and

Table 7: Education, productivity, wages and employment in the quantitative model

	Base		Human cap ($\gamma \times 2$)		Signal ($\sigma = 0$)		Lower costs	
	With SE	No SE	With SE	No SE	With SE	No SE	With SE	No SE
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Proportion with educ.	0.30	0.40	0.37	0.48	0.57	0.57	0.59	0.71
High skilled educated	0.73	0.75	0.67	0.74	1.00	1.00	0.69	0.69
High skilled uneducated	0.50	0.45	0.52	0.42	0.00	0.00	0.40	0.29
Proportion of self-emp.	0.34	0.00	0.28	0.00	0.08	0.00	0.33	0.00
$H(\hat{\phi}_S)$	0.18	0.00	0.13	0.00	0.14	0.00	0.36	0.00
$H(\hat{\phi}_N)$	0.41	0.00	0.37	0.00	0.00	0.00	0.28	0.00
Prod per employee	4.22	4.15	5.18	5.40	6.01	6.07	4.54	6.14
e=S	5.76	6.30	7.45	8.18	7.45	7.45	5.07	7.08
e=N	3.49	2.76	3.39	2.93	4.74	4.74	3.83	3.50
Prod self-employment	4.55		4.71		5.95		4.93	
e=S	5.66		5.95		5.95		5.09	
e=N	4.34		4.46				4.63	
Per capita prod	4.04	3.41	4.61	4.48	4.91	5.09	4.40	5.10
e=S	4.68	5.20	6.45	6.74	5.38	5.73	4.85	5.96
e=N	3.76	2.21	3.54	2.42	4.27	4.22	3.75	2.96
Mean wage								
e=S	4.90	4.62	5.91	5.70	5.31	4.93	3.90	4.48
e=N	2.90	2.17	2.82	2.34	3.97	3.91	3.23	2.80
Labor mkt income								
e=S	4.67	4.14	5.33	4.83	5.39	4.51	3.53	3.90
e=N	2.67	2.01	2.69	2.14	3.37	3.30	2.96	2.71
Labor income								
e=S	4.62	4.19	5.14	4.81	5.91	4.91	3.55	4.04
e=N	2.75	2.13	2.77	2.31	3.02	2.98	3.05	3.05
No wage offer	0.111	0.191	0.114	0.176	0.148	0.179	0.107	0.157
e=S	0.083	0.175	0.098	0.175	0.183	0.230	0.109	0.159
e=N	0.124	0.202	0.124	0.176	0.100	0.110	0.104	0.154
Unemployment	0.087	0.191	0.096	0.176	0.141	0.179	0.089	0.157
e=S	0.078	0.175	0.093	0.175	0.168	0.230	0.086	0.159
e=N	0.088	0.202	0.095	0.176	0.100	0.110	0.092	0.154
Employed (%)								
BN	0.68	0.61	0.56	0.53	0.53	0.51	0.43	0.26
BS	0.00	0.00	0.02	0.01	0.00	0.00	0.07	0.15
WN	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
WS	0.32	0.39	0.42	0.46	0.47	0.49	0.50	0.59

then examining the effects these parameter changes have on worker allocation.

5.1 Higher human capital accumulation

We turn now to considering how these outcomes are affected by changes in parameters that characterize the quality of education.

First, we double the accumulation of human capital through education ($\gamma \simeq 1$ instead of $\gamma \simeq 0.5$) to represent an education with higher value added to students. We report the results in columns (3) and (4) of table 7. This change impacts education decisions significantly: the share of educated workers rises to 37%, mostly due to an increase in education among those with low skills.

Productivity, then, increases substantially from 4.22 to 5.18, and per capita productivity rises

from 4 to 4.6. This higher productivity is explained by several effects. There is, first, a direct effect in productivity of the educated employees, which goes from 5.76 to 7.45. Second, some of the unskilled workers decide to invest in education, increasing their productivity. Third, the higher job productivity raises vacancy creation, particularly for type- W firms. Fourth, higher wages in the labor market reduce incentives to self-employ, thereby inducing a rise of self-employment productivity from 4.55 to 5.71. This is because only workers with the best idiosyncratic shocks choose self-employment. Thus, the effect of higher human capital accumulation is amplified through worker allocation.

In column (4) of table 7 we report the results of the same increase in human capital accumulation without self-employment, comparable to the economy reported in column (2). Education grows to 48% and productivity per employee goes up to 5.4, while per capita productivity reaches 4.48. It is important to emphasize the one mechanism generating this effect, in addition to the ones described above. Without self-employment, workers are more induced to educate and, when participating in the labor market, they allocate in very productive white-collar jobs (that yield more than 8 units of product) instead of working as self-employed (earning a mean of 5.9 units of product).

5.2 Better quality of the signal

Consider the case where there are no idiosyncratic costs of education. This means that, since $\sigma = 0$, all workers of a given type have the same cost of education.

Column (5) of table 7 reports these results in an economy with self-employment. The proportion of educated workers rises to 57% and education provides a perfect signal, with a separating equilibrium: all highly skilled workers invest in education, while all those with low skills avoid education.

This separating equilibrium also implies that high-skill workers are exclusively employed in the job type where they are most productive, i.e. in white-collar jobs. There are no unskilled workers employed by white-collar firms. This allows for an increase in productivity from 5.76 to 7.45 for educated (and skilled) employees, and from 3.49 to 4.74 for uneducated employees. It is worth noting that this increase in productivity is exclusively driven by a better allocation of skills, given the signaling role of education, since this separating equilibrium allows firms to employ just those workers that are most productive in their jobs. Mean wages for both educated and uneducated workers go up, even when unemployment is now higher.

A crucial source of higher productivity through allocation is the smaller proportion of self-employed workers. With appropriate signaling and the rise in wages, workers are less prone to work as self-employed. For that reason, the share of self-employment goes down to 8% compared to 34% in the baseline economy. Moreover, all the self-employed are now high-skill workers, so no unskilled worker decides to self-employ. This provides further incentives to educate through equation (7).

Importantly, even when the highest productivity for unskilled workers in self-employment (3.19) is lower than the productivity as employee (4.74) unskilled workers decided to self-employ because their wage was low.

5.3 Lower costs of education

When education entails human capital accumulation, a reduction in educational costs could be justified as a means to distribute the benefits of higher human capital to all workers. Nevertheless,

this objective could be at odds with the signaling motive of education, and the outcome of such a policy is unclear.

We then use our model to analyze the effect of a reduction in the costs of education. We first consider a cost reduction of the same size for all workers, changing our cost function to $\hat{c}(S, \theta, \varepsilon) = c(S, \theta, \varepsilon) - c_0/2$. We report the results in columns (7) and (8). Education increases to 59%, but at the expense of pooling: 30% of educated workers are unskilled and 40% of uneducated workers are skilled. On the one hand, this implies an increase in average productivity of workers. On the other hand, there is an allocation problem, because firms demanding workers among the educated (i.e. type- W firms) are more likely to be matched with unskilled workers, with much lower productivity. These effects compensate each other given that the average productivity of employees increases only mildly. There is a reduction in the productivity of the educated and an increase in the productivity of the uneducated. Per capita productivity is now slightly higher.

Given the lower educational costs, more educated workers decide to self-employ while less uneducated workers follow that path. There is a sharp difference in allocation in the labor market: 7% of all employees work in blue-collar jobs even when they are educated. There are not enough jobs for educated workers as white collar, partly because their productivity has gone down due to the larger number of unskilled workers that are mixed within the educated group. The loss in the signaling quality of education thus worsens worker allocation, since highly skilled workers could otherwise be employed in high-productivity, white-collar jobs.

On the whole, we find that a reduction in the costs of education improves welfare of workers as a whole, as measured through expected labor income. Nevertheless, this improvement is restricted to the unskilled workers that take advantage of the reduction in the costs of education. These gains are at the expense of welfare losses of the skilled workers for which expected labor income falls.

In column (8) of table 7 we report the result of changing the costs of education in an economy without self-employment. As in other cases, absent the self-employment option, incentives to educate grow, particularly for low-skill workers. The proportion of educated workers goes up to 71%, with many unskilled workers investing in education.

6 Conclusions

Through the model and its quantitative implementation we find that there is an important interaction between human capital accumulation, the signaling quality of education, labor market performance and self-employment. In particular, we find that a lower quality of education increases self-employment, under the condition that education does not increase self-employment productivity directly. Additionally, any negative effect of lowering the quality of education (both through lower human capital accumulation or through a less precise signaling role) can be amplified through self-employment. The interaction between human capital accumulation and signaling is particularly important. Shocks to human capital can be amplified when signaling quality is not perfect, since partial pooling could arise and enhance the direct effects of lower productivity.

We also explore whether the productivity (and income) improvements of reducing costs of education could be offset by the worsening of the signaling problem and the partial pooling that could arise. We found that this could be the case. The reason is that partial pooling reduces productivity through misallocation of abilities, given that highly skilled workers cannot be identified through education.

On the whole, with lower quality of education workers tend to reduce their investment in education, lowering skills and affecting the productivity of the self-employed.

References

- [1] Alvarez, F. and M. Veracierto, "Labor-Market Policies in an Equilibrium Search Model," *NBER Macroeconomics Annual*, 1999, Volume 14, National Bureau of Economic Research, 2000.
- [2] Arozamena, L. and H. Ruffo (2015) "Signaling and Misallocation: Education and Labor Markets", mimeo.
- [3] Banerjee, A. V. and A. F. Newman (1993). Occupational Choice and the Process of Development. *Journal of Political Economy* 101(2), 274–98.
- [4] Buera, F. J., J. P. Kaboski, and Y. Shin (2011). Finance and Development: A Tale of Two Sectors. *American Economic Review* 101(5), 1964–2002.
- [5] Burdett, Kenneth and Mortensen, Dale T (1998). Wage Differentials, Employer Size, and Unemployment. *International Economic Review* 39(2), 257-73.
- [6] D’Erasmus, P. N. and H. J. Moscoso Boedo (2012). Financial structure, informality and development. *Journal of Monetary Economics* 59(3), 286–302.
- [7] Erosa, A. and L. Allub (2012). Financial Frictions, Occupational Choice, and Economic Inequality. 2012 Meeting Papers 702, Society for Economic Dynamics.
- [8] Heckman, J. and T. Kautz, "Hard Evidence on Soft Skills," *Labour Economics*, 19, 451-464, 2012.
- [9] Hopenhayn, H. A. (2014). Firms, misallocation, and aggregate productivity: A review. *Annual Review of Economics* 6(1), 735–770.
- [10] Mano, Y., A. Iddrisu, Y. Yoshino, and T. Sonobe (2012). How Can Micro and Small Enterprises in Sub-Saharan Africa Become More Productive? The Impacts of Experimental Basic Managerial Training. *World Development* 40(3), 458–468.
- [11] Marchionni, M. and E. Vázquez (2015). The effect of schooling on skills and knowledge in Latin America: evidence from PISA. Mimeo.
- [12] Midrigan, V. and D. Y. Xu (2014). Finance and Misallocation: Evidence from Plant-Level Data. *American Economic Review* 104(2), 422–58.
- [13] Mortensen, D. (2005) *Wage Dispersion: Why Are Similar Workers Paid Differently*. The MIT Press
- [14] Shimer, R., "Mismatch," *American Economic Review*, 97, 1074-1101, 2007.
- [15] Spence, M. *Market Signalling: Information Transfer in Hiring and Related Processes*, Harvard University Press, 1973.

A Appendix

A.1 Description of data sources

In this section we present a brief description of the data that we use in the main text. These are the Survey of Adult Skills (PIAAC) which covers OECD countries; the Skills Measurement Program (STEP) from the World Bank with data for many developing countries, including Bolivia and Colombia; and the *Encuesta Nacional de Habilidades y Mercado Laboral* (ENHAB), carried out by the World Bank in Peru.

These sources have the same main objective, which is to assess the skills of labor supply in each country. The aim is ambitious in that it implies measuring the whole distribution of a vector of skills for the adult population. Additionally, it requires including broad contextual information to link skills to education and the labor market.

A main concern with these sources is the how comparable their indicators are. PIAAC and STEP are implemented in different countries, languages and contexts. In both surveys, special care is taken so that comparability is ensured. Each step of the process (from translation of instruments to scoring) was designed to allow for comparable measures. One piece of evidence about this are the results in the cross-country reliability studies implemented in PIAAC. There, bilingual scorers were asked to score both national and international items. The agreement in scoring was above 95% in all countries, with the exception of France. In addition, the design of the STEP survey made explicit efforts to provide scores of literacy skills that are comparable to those in PIAAC. This gives the opportunity to contrast a set of developing and OECD countries using the same scale and dimension of cognitive skills.

In this type of surveys, skills are both cognitive and socio-emotional, recognizing that both kinds of skills can be important in a job (manual dexterity, potentially important as well, is not measured in these surveys; see Prada and Urzua, 2014). Cognitive skills include literacy, vocabulary, numeracy, problem solving, and working memory, among others. Socio-emotional skills include the so-called Big Five personality trait factors (conscientiousness, openness, neuroticism, agreeableness, extraversion), grit (perseverance for long-term goals) and risk aversion.

Each survey that we use assesses a subset of these skills. PIAAC measures cognitive skills, namely literacy, numeracy and problem solving, but does not include any socio-emotional ability. STEP measures literacy and a variety of socio-emotional skills (Big Five, Grit and risk aversion). Finally, ENHAB evaluates vocabulary, verbal fluency, working memory and mathematics problem-solving, all of them cognitive skills, while it assess the Big Five personality traits and grit as well. As emphasized above, literacy measures are comparable in PIAAC and in STEP.

Additionally, all these surveys have background and contextual information on the respondents. They gather detailed information on education and training, with the objective of understanding how these skills are affected by them. They also provide information on labor market history, current work status, current occupation and labor income, and skills use at work.

In what follows, we briefly describe each of these sources.

A.1.1 PIAAC

The OECD Survey of Adult Skills (PIAAC) is intended to measure skills that are essential for participating in societies, and thereby derive distributions of the adult population each of these skills. The skills assessed in the survey are those related to information-processing. The survey covers all residents aged 16-65 within each country and was implemented between 2011 and 2012

(Round 1) and 2014 to 2015 (Round 2).⁸ The required sample size for each country is of 5000 completed cases, and the survey has two main components: contextual information and cognitive assessment.

The main skills measured in the PIAAC surveys are literacy, numeracy and problem solving. As mentioned above, they are considered to be “key information-processing skills,” since they are crucial for accessing, understanding, analyzing and using information. They are also necessary for fully integrating and participating in the labor market, education and training, and social and civic life. It must be stressed that these skills can be acquired in different contexts: in formal education, in training and with experience. Additionally, literacy and numeracy are required for developing new cognitive skills, given that they provide understanding for gaining access to new information.

Specifically, literacy is taken as “the ability to understand, evaluate, use and engage with written texts”. It includes the assessment of the decoding of written words and sentences, as well as the comprehension, interpretation, and evaluation of complex texts. Numeracy is “the ability to access, use, interpret and communicate mathematical information and ideas”. It includes the assessment of “managing a situation or solving a problem in a real context, by responding to mathematical content/information/ideas represented in multiple ways”. Problem solving is described as “the ability to use digital technology, communication tools and networks to acquire and evaluate information, communicate with others and perform practical tasks”. The assessment focuses on the abilities to solve problems for personal, work and civic purposes by setting up appropriate goals and plans, and accessing and making use of information through computers and computer networks.

PIAAC was administered in the respondent’s home, and the cognitive assessment was carried out in one session. About 80% of the respondents took a computer-based format instrument while the remaining 20% was paper-based. All individuals were required to take both literacy and numeracy items; problem-solving items and reading components were taken by a subset of individuals. Additionally, some countries (such as Finland, France) did not apply reading components, while France, Italy and Spain did not implement problem-solving items. Scoring was automatic in the computer-based format. Manual scoring was necessary in the case of respondents taking the paper-based version. A set of within country and cross-country reliability studies were implemented, producing an agreement in scores higher than 95% both within and between countries. The only exception is France, with 87% of agreement in literacy scores.

Besides the instruments for measuring cognitive skills, PIAAC also provides important contextual information of each respondent, including education, work status and variables about labor market history, occupation, among others. Additionally, it includes details about literacy and numeracy practices and their use at work. It also gathers information about the extent to which individuals are required to use a range of skills in their work. Respondents also report on how and whether their skills and qualifications match those required for their jobs. This last type of information has been used to examine the mismatch of abilities and skill requirements in jobs (see, for example, Pellizzari and Fichen, 2013, and Adalet McGowan and Andrews, 2015).

A.1.2 STEP

The World Bank’s Skills Towards Employability and Productivity (STEP) program aims at understanding the how skills relate to employability and productivity (Pierre, Sanchez-Puerta, Valerio,

⁸PIAAC covers Austria, Belgium, Canada, Denmark, Estonia, Finland, France, Great Britain, Germany, Ireland, Italy, Japan, S.Korea, Netherlands, Norway, Poland, Russia, Slovakia, Spain, Sweden and the US. See OECD (2012a, 2012b,2013)

Rajadel, 2014). For that purpose its goal is to measure human capital stocks, as skills supply. It provides an assessment of skills in the adult population and the use of these skills in jobs. It covers several developing countries between 2012 and 2014. We use information of Bolivia and Colombia, with about 2,500 observations in each country.

The STEP program developed survey instruments tailored to collect data on skills in low- and middle-income country contexts. It is implemented as a household survey. The instruments to assess skills are administered on adult respondents aged between 15 and 64 selected at random within each household.

The definition of skills in STEP includes cognitive, socio-emotional and job-relevant skills. The assessed cognitive measure is reading literacy. A main goal of the design of the literacy assessment is providing a direct link to the literacy scale used in PIAAC (World Bank, 2014). For that purpose, STEP implements a 45-minute assessment that allows countries to be on the same reading literacy scale as the one used in PIAAC. Additionally, to provide information about the full distribution of literacy skills in the adult population in each country, a main goal of the literacy assessment design is correctly measuring the lower end of the literacy scales.

Socio-emotional skills are measured through a series of items, related to the Big Five personality trait factors (conscientiousness, openness, neuroticism, agreeableness, extraversion). Each of these factors is assessed with three items in the short Big Five Inventory. Additionally, grit (perseverance for long-term goals) and risk aversion are measured independently. Aggregate measures by trait are constructed by STEP. So as to adequately measure human capital supply, STEP provides background information on households and detailed information about individuals within the household, including: skills acquisition history, educational attainment, work status and labor market history, family background, and health.

A.1.3 ENHAB

The main objective of ENHAB (“Encuesta de hogares sobre habilidades y funcionamiento del mercado laboral peruano”) is measuring workers’ skills and their labor market history. In particular, it focuses on each respondent’s first job.⁹

ENHAB was developed by the World Bank and implemented by “Instituto Cuánto”. It was carried out in of 2010 in 17 Peruvian cities. It is a nationally representative household survey comprising information of urban households. Within each household, one adult between 18 to 50 years of age is randomly selected to participate in the skills assessment. The sample comprises 2666 adults.

Skills is defined broadly, including cognitive and socio-emotional skills. Cognitive skills are measured through Peabody Picture Vocabulary Test (fourth edition), verbal ability (phonemic verbal fluency), working memory (digit memory test) and mathematics problem-solving. Socio-emotional skills are evaluated adapting standardized methods related to the Big Five factors. Each factor is assessed through seven items, with self-assessment responses. Additionally, the survey measures grit through 17 self-assessment items. These methods are comparable to those applied in STEP.

A set of general questions and household background information are also available in this survey. In particular, household living conditions, demographic information, academic achievement, current employment, occupation and earnings, among others, are relevant for our purposes.

⁹See Yamada et al. (2014), Abusada et al. (2015), and Diaz et al. (2012).

A.2 Results for calibration

For calibration we use ENHAB. We restrict the sample to individuals aged between 25 and 40 that participate in the labor market. We construct binary variables on education (that identifies workers with education higher than completed tertiary or college degree), occupations (that identifies technicians or professionals) and skills. The skills variable is constructed as follows. First we consider all cognitive and socio-emotional skills in the survey. Following the idea of latent variables (see Heckman et al., 2006) we include these variables to extract a single factor running an explanatory factor analysis. A prediction of the factor levels using factor loadings generates a continuous variable that we consider as skills. Using this skills measure, we divide the respondents between low- and high- skill setting a cutoff value close to the median of the factor in the sample. We compute the proportion of workers according to the binary education, skills and occupation variables that were used as targets of the model.

The calibration procedure included measures of wage and labor income differentials. For that purpose we regressed labor income on our binary variables. Table A.1 shows the results of OLS regressions of the log of monthly labor income on the education binary variable (column 1), on skills (column 2) and on occupation (column 3), restricting the sample to wage earners. Additionally, in columns 4 and 5 we show the result of OLS regressions of the log of self-employed monthly labor income on education and skills. These coefficients represent wage differentials in log points. We transform these to wage ratios by applying exponentials to the coefficients and we use these as targets for calibration.

Table A.1: Regression of the log of labor income

	(1)	(2)	(3)	(4)	(5)
Education	.4081*** (0.0971)			.3261** (0.155)	.2076 (0.1570)
Skills		.2804*** (0.1049)			.4340*** (0.1232)
Occupation			.3067*** (0.1039)		
N	330	330	330	320	320

Notes: . Sample is restricted to employed wage earners (columns 1 to 3) or self employed (columns 4 to 5) from 25 to 40 years of age. *, **, and *** indicate significance at the 10%, 5% and 1% levels.
Source: ENHAB.

A.3 Model results for calibration

In this appendix describe how we use elasticities and unemployment levels to calibrate the vacancy cost function, $\kappa(v) = av^\alpha$. We consider a single labor market with n workers and n firms. Workers have an option flow of b , while firm productivity for a filled job is p . Each firm can open v vacancies

at a cost, given by $\kappa(v)$. When opening vacancies, firms offer a fixed wage, taking advantage of monopsony power.

When productivity increases, firms have higher incentives to open vacancies. It can be shown that the elasticity of unemployment with respect to productivity is

$$\begin{aligned}\eta_{u,p} &= \frac{du/u}{dp/p} = -\frac{p}{(p-b)} \left(\frac{1}{1 + \frac{n}{m} \frac{c''}{c'}} \right) \\ &= -\frac{p}{(p-b)} \left(\frac{1}{1 + \frac{\alpha-1}{\lambda}} \right)\end{aligned}$$

and thus

$$\alpha = 1 + \lambda \left(\frac{1}{|\eta_{u,p}|} \frac{p}{p-b} - 1 \right)$$

The value of λ determines unemployment in the model: $u = \exp(-\lambda)$ in a model with one market. Setting unemployment to the aggregate value of 10% yields $\lambda = -\ln(0.1) = 2.3$. Using $p = 4b$, and the observed $\eta_{u,p} = -0.55$, we get $\alpha = 4.3$. Additionally, using the first-order condition for vacancy creation, and following analogous steps, we get $a = 4 \times 10^{-5}$. We initialize our calibration procedures with these values.

A.3.1 The effect of the number of workers on unemployment

To understand the effect of changes in the number of workers in a market it is useful to derive the elasticity of the number of vacancies with respect to the number of workers, n . If this elasticity is lower than one, then an increase in the number of workers would increase unemployment.

We illustrate this effect using a single market. Taking the CPO of the firm, $e^{-\lambda}(p-b) = c'(v)$, and differentiating with respect to n and v we get

$$\frac{dv}{dn} \frac{n}{v} = \frac{1}{1 + \frac{c''}{c'}}.$$

Thus, more workers, n , into a market with given firms, m , would generate higher unemployment depending on the convexity of the cost function. If the cost function is linear, the elasticity would be 1 and unemployment would not change. If the cost function is strictly convex the elasticity would be strictly lower than one and unemployment would rise with the number of workers.

The intuition is that convex cost functions imply decreasing returns to scale in the matching process so that increasing the number of workers would imply higher marginal costs to firms. In such a case, the firms can take advantage of their monopsony power and increase profits.

The convexity of the vacancy cost function depends on the calibrated parameter α , as well as on equilibrium vacancies. In particular, $\frac{c''}{c'} = \frac{(\alpha-1)}{v}$. When we consider approximations to these values according to the baseline economy, we get $\frac{3.3}{10} = 0.33$. Then $\frac{dv}{dn} \frac{n}{v} = \frac{1}{1.33} \simeq 0.75$. This implies that an increase in 10% of the number of new workers in a market would lead to an increase of 7.5% in the number of vacancies, thereby raising unemployment, as $\frac{mv}{n}$ goes down by about 2.3%.

References

- [1] Abusada, E., P. Lavado and G. Yamada (2015) “Skills for first job”. U del Pacifico. Mimeo.
- [2] Adalet McGowan, Muge and Dan Andrews (2015) “Labour Market Mismatch and Labour Productivity: Evidence from PIAAC Data”. OECD Economics Department Working Papers Nro 1209.
- [3] Diaz, J. J., O. Arias and D. Vera Tudela (2012) “Does perseverance pay as much as being smart? The returns to cognitive and non-cognitive skills in urban Peru”. Mimeo.
- [4] Heckman, J., J. Stixrud and S. Urzua (2006) “The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior”, *Journal of Labor Economics*, 24, 3, 411-482, 2006.
- [5] OECD (2012a), *Literacy, Numeracy and Problem Solving in Technology-Rich Environments: Framework for the OECD Survey of Adult Skills*, OECD Publishing.
- [6] OECD (2012b), *The Survey of Adult Skills. Readers companion*, OECD Publishing.
- [7] OECD (2013), *Technical Report of the Survey of Adult Skills*, OECD Publishing.
- [8] Pellizzari, M. and A. Fichen (2013) “A new measure of Skills Mismatch”. OECD Social, Employment and Migration Working Papers Nro 153.
- [9] Prada, M. F. and S. Urzúa “One Size does not Fit All: Multiple Dimensions of Ability, College Attendance and Wages”, NBER Working Paper 20752, National Bureau of Economic Research, 2014.
- [10] Pierre, G., M. L. Sánchez-Puerta, A. Valerio and T. Rajadel (2014), “STEP Skills Measurement Surveys. Innovative Tools for Assessing Skills”. The World Bank. July 2014.
- [11] World Bank (2014) “A guide to understanding the literacy assessment of the STEP skills measurement survey”. November 2014.
- [12] Yamada, G., P. Lavado, and L. Velarde (2014), “Habilidades no cognitivas y brecha de género salarial en el Perú,” *Estudios Públicos*, 0, 89129, 2014.