

Firm-Provided Training, Selection and Education

Marco A. Barrenechea-Méndez, PhD.[†]

Universidad Pública de Navarra

E-mail: Marco.barrenechea@unavarra.es

Telephone number: 0034699617540

Pedro Ortín-Ángel, PhD.

Universitat Autònoma de Barcelona

E-mail: Pere.ortin@uab.es

Telephone number: 0034935811209

Eduardo C. Rodes, PhD.

Universitat Autònoma de Barcelona

E-mail: Eduardo.rodes@uab.es

Telephone number: 0034935811209

Objectives: The objective of this paper is to provide theoretical and empirical analyses of the joint effects of the requirements (e.g. education credentials) and intensity (e.g. screening) of the firm selection process on firm-provided training. The economic literature has analyzed, independently, the relation between training and education credentials and between training and selection intensity. This paper integrates this two streams of literature into a unique theoretical framework

Methodology: we provide a theoretical model as well as empirical evidence. We use a database containing information of firm organizational practices of Spanish industrial establishments. We utilize a series of probabilistic models (OLS, Probit and TSLS) in order to test the predictions of the theoretical model.

[†] The authors received financial support from the Spanish Ministry of Economy, Industry and Competitiveness, R&D project: ECO 2013-48496-C4-4-R.

Results: Our evidence shows that the positive relationship between firm-provided training and education generally found in the literature is, in our case, indirect. The variability of both variables training and education are explained by the firms' selection intensity.

Abstract: We provide theoretical and empirical analyses of the relationship between firm-provided training and selection policies. We focus on two features of the selection process: requirements (education level) and intensity (information collected). Our evidence shows that the positive relationship between firm-provided training and education generally found in the literature is, in our case, an indirect one. The variability of both variables are explained by the firms' selection intensity. The theoretical model argues that firms with more intense information collection processes will be more confident about the quality of this information and thus will set higher selection requirements and provide more training.

Key words: Training, Selection and Education.

JEL: M53, M54, M12, J24

Presently, it is known that firm-provided training plays an important role in workers' skill development. As a matter of fact, among the 36,000 workers surveyed by the European Working Conditions Survey 2015, approximately 36 percent reported that they underwent training paid for or provided by their employers¹. The analysis of the determinants of this type of training has received considerable attention in the economic literature. This paper is related to the stream of economic literature analyzing the relationship between firms' investments in training and their workers' selection policies.

Following Barron, Black, and Loewenstein (1989) and Barron, Bergen, and Black (1997), we refer to firms' selection policies as the process of collecting information about workers' abilities in order to decide on their hiring. In this process, firms can use several sources of information, such as credentials on education, ability tests, interviews and the

¹ In this paper, we do not make a distinction between firm-specific and general human capital. Moreover, in the theoretical model, we assume that only the firms pay the costs of a training investment. This assumption is consistent with the observation that firms usually pay a share of training costs even though the skills provided are largely general (Pischke 2001).

like. Hence, all this information can be summarized in the estimated abilities of the workers and the precision of these estimates. Therefore, firms' selection policies can be described according to two main features: i) their requirements² (the minimum estimated ability required to be employed) and ii) their intensity (the precision of the abilities' estimations).

The analysis of the relation between firm-provided training and selection requirements has generally focused on the role of education credentials (Spence, 1973) as a source of information on workers' abilities. The relation between firms' training investments and workers' educational attainment has been extensively studied (Rosen, 1976, Mincer, 1992, Heckman, 2000), and the related empirical evidence generally shows that these variables are positively associated. This result is almost a tenet in economics, it holds practically without exception across different datasets, groups of workers and countries; the argument behind this result is the complementarities between these two variables.

The existing analysis of the relation between firm-provided training and selection intensity is less extensive. Barron, Black, and Loewenstein (1989) and Barron, Bergen, and Black (1997) find a positive association between selection intensity and firm-provided training³. Their explanation for this finding relies on the need for a good match between firms' requirements and workers' abilities when the importance of the training investment increases.

In this paper, we provide theoretical and empirical analyses of the joint effects of selection requirements and intensity on firm-provided training. In that regard, the paper can be considered an effort to integrate the two streams of the literature cited above. As far as we know, no other previous efforts have aimed to provide such analysis.

² We can justify the presence of these requirements based on standard human resources practices. A basic tool for recruitment and selection is the use of a job description. Typically, the job description is a short document that describes the tasks, duties and responsibilities of a job position. It always contains the job specifications and minimum skill requirements (Mathis et al. 2014, p.139).

³ There is a stream of literature (Stevens 1994; Lindley 1975; Merrilees 1980; Mason et al. 2012) providing evidence of a negative relationship between recruitment costs and firm-provided training. Therefore, if recruitment costs and selection intensity are positively correlated, this evidence can be understood as contrary to the evidence cited above.

From a theoretical point of view, the paper analyzes a context in which firms have different selection intensities. Based on the selection intensity, the firm simultaneously establishes the selection requirements and training level associated with the job position and then looks for a candidate. In this situation, differences in firms' selection intensities are explaining, at the same time, differences in selection requirements and training. When firms have better information about candidates, they will set higher requirements in the selection process and will provide more training to their workers.

A consequence derived from the model is that the standard positive relationship between firms' investments in employees' training and the degree of schooling of those employees will be reduced when we control for a measure of selection intensity. The reason for this reduction is that the selection requirements and intensity are expected to be positively correlated.

To test these implications, we collect a database containing information on blue-collar workers' human resources practices in Spanish industrial plants. Typically, measures of firms' training investments, firms' selection intensity and employees' education attainment are not available in the same dataset. For instance, we have not found such information in surveys such as the European Company Survey (2004, 2009 and 2013), the European Working Conditions Survey (1991, 1996, 2000, 2005, 2010 and 2015) or the Workplace Employment Relations Study (2011). The evidence provided strongly supports the implications of the model.

The remainder of the paper is organized as follows. In the next section, we review the relevant literature more extensively. In section II, we lay out the theoretical benchmark. Data and empirical methods are presented in section III, and the results are given in section IV. In section V, we conclude by drawing conclusions and implications of the paper.

I. Literature Review

The association between firm-provided training and workers' education attainment has been investigated largely across different datasets, groups of workers and countries, such as the US (Lillard, and Tan 1986; Brown 1989; Tan et al. 1992; Altonji and Spletzer

1991; Royalty 1996; Lynch and Black 1998; Barron, Black, and Loewenstein 1989; Lynch 1992), the UK (Greenhalgh and Stewart 1987; Booth 1991, 1993; Tan 1991; Green 1993; Greenhalgh and Mavrotas 1994; Blundell et al. 1996; Arulampalam, Booth, and Elias 1996; Arulampalam and Booth 1998), Germany (Pannenberg 1996; Pischke 2001; Kuckulenz and Zwick 2003), the Netherlands (Groot et al. 1994; Oosterbeek 1996 and 1998), Australia (Tan 1991), France (Goux and Maurin 2000) and a group of 13 European Union countries (Brunello 2001). Scholars have used different measures to estimate different specifications of the following equation:

$$(1) \quad t_i = \alpha_1 + \beta_1 e_i + \epsilon_{i,1}$$

where t is a measure of firm i 's training investment, e a measure of workers' investments in education, and $\epsilon_{i,1}$ is the error term. Parameters α_1 and β_1 are to be estimated. Almost without exception⁴, this empirical literature provides evidence of a positive association between training and education, $\beta_1 > 0$.

Equation (1) is generally interpreted in these empirical applications as the first-order condition of the maximization problem that firms solve to decide workers' optimal level of training. For illustrative purposes, we can write this problem as the maximization of the following objective function, that is, the benefits of education per worker π :

$$\max_t \pi(t) = at - zt^2/2 + z\beta et + tz\epsilon$$

From the first-order condition (FOC: $a - zt + z\beta e + z\epsilon = 0$), one can deduce Equation (1):

$$t^* = \alpha + \beta e + \epsilon,$$

where β is a measure of the complementarities between training and education, $\alpha = \frac{a}{z}$ and ϵ is the random shock. To obtain an optimum, it is necessary that the profit margin

⁴ We are aware of few exceptions. For example, Booth (1993) found that the probability of men receiving this type of training is higher for first-class degree holders; however it is less prevalent among people who have subsequently pursued postgraduate education. Among women, first-class degree holders receive less training than other types of graduates. Tan et al. (1992) found that in general, the probability of receiving most kinds of training rises with the level of education attainment, though unlike for the US and the UK, the evidence for Australia is quite weak. Blundell et al. (1996) found that the likelihood of receiving employee-provided training grows with educational qualifications. However, post-school qualifications have no significant effect on the probability of obtaining employer-provided training.

decreases with the level of training, $z > 0$. Parameter $a > 0$ assures the existence of positive profit margins when there are no complementarities ($\beta = 0$). In empirical implementations, coefficients α_1 and β_1 provide the estimates of α and β , respectively.

Note that from the estimation of Equation (1), we cannot make inferences about the possible sources of the complementarities between training and education. Considering that the objective function above results from comparing the productivity of training with its costs: $\pi(t) = P(t) - C(t) = at - \frac{zt^2}{2} + z\beta et + tze$, parameter β could be interpreted as a reduction in costs (i.e., $P(t) = at$ and $C(t) = \frac{zt^2}{2} - z\beta et$). This interpretation is suggested, for example, by Rosen 1976; Blundell et al. 1996; Arulampalam and Booth 1998; Oosterbeek 1998; Heckman 1999; Pischke 2001. Alternatively, β can be interpreted as an increase in the benefits of training investment (i.e., $P(t) = at + z\beta et$ and $C(t) = \frac{zt^2}{2}$), as suggested by Arulampalam and Booth 1998; Oosterbeek 1998; Heckman 1999 and Pischke 2001.

Furthermore, from Equation (1), we cannot infer whether the association across plants between workers' average training and their average educational attainment in a similar job position is a direct or an indirect one. As a direct relationship, we refer to a situation where plants hire workers and adapt the training level provided to the characteristics of the workers hired. In this case, a positive relationship between training and education is expected when these two variables are complementary and different plants hire workers with different levels of education. As an indirect relationship, we refer to a situation where firms establish the training level and the abilities or personal characteristics of the workers that they want for the job position, and then, they look for candidates in the labor market. Therefore, the variability across plants of the relation between training and education can be explained by an exogenous omitted variable.

Another stream of literature (Barron, Black, and Loewenstein 1989; Barron, Bergen, and Black 1997) has analyzed the relationship between training and the workers' selection intensity. The results of those analyses can be summarized by the estimation of the following relation:

$$(2) \quad t_i = \alpha_2 + \delta_2 \rho_i + \epsilon_{i,2}$$

where ρ_i is a measure of selection intensity, which in the cited works is positively related with the firms' investments in training ($\delta_2 > 0$). Together, the two streams of literature previously analyzed suggest that firms maximize the following problem:

$$\max_t \pi(t) = at - \frac{zt^2}{2} + z\beta et + z\delta\rho t + tz\epsilon$$

From the first-order condition (FOC: $a - zt + z\beta e + z\delta\rho + z\epsilon = 0$), the following equation can be deduced:

$$t^* = \alpha + \beta e + \delta\rho + \epsilon$$

considering that $\alpha = \frac{a}{z}$. Barron, Berger, and Black (1997)'s empirical analysis relied on four datasets of American employers and found strong support for the prediction that employers using more intensive selection processes employ workers with higher levels of education. These results suggest that the investment in selection is positively related to the years of education of the employees selected. Assuming that this relationship is established in the following terms, $e = \lambda\rho$, given $\lambda \geq 0$, the expected value of parameter β_1 can be expressed as $\beta + \delta/\lambda$, i.e., as the sum of direct and indirect effects. In other words, the relationship between training and education could be attributed to omitted relevant variables related to the selection process.

Our research proposal is to test whether selection intensity is one of these relevant omitted variables. First, we elaborate a simple theoretical model explaining why selection intensity could positively affect the level of training ($\delta > 0$) and the levels of education of the workers selected ($\lambda > 0$) and why the relation between firms' investments in training and their workers education attainment could be (partially or totally) indirect. Second, we provide empirical evidence for the predictions of the model.

The empirical application is consistent with the firms' usual sequence of decisions. First, firms evaluate the candidates available with different levels of intensity. After the evaluation process, only the selected ones join the firm, determining the average educational level of firms' staff. Finally, firms provide training to the employees selected. Considering that an effect cannot occur from a cause which is not in the past, the paper postulates estimating Equations (3) and (4) where the dependent variables in the empirical

analyses are going to be the average educational level of the firm's staff and the firm provided training:

$$(3) \quad e_i = \alpha_3 + \lambda_3 \rho_i + \epsilon_{i,3}$$

$$(4) \quad t_i = \alpha_4 + \beta_4 e_i + \delta_4 \rho_i + \epsilon_{i,4}$$

where the expected value of β_4 is β , so we can test whether the effect of education on training is totally indirect, ($\beta_4 = 0$, $\lambda_3 > 0$ and $\delta_4 > 0$), or there is some direct effects ($\beta_4 > 0$). Take note that Equations (1) and (2) are particular cases of Equation (4) [Equation (1) imposes $\delta_4 = 0$ and Equation (2) $\beta_4 = 0$].

II. Theoretical Framework

A. The general framework: selection procedure and timing

We set up a general framework to analyze the provision of training when the selection process provides firms with information about the candidates' abilities. We are interested in the empirical implications of the model regarding selection intensity, education and training levels – in other words, on the coefficients of Equation (3) and (4).

Period 1. There is a pool of heterogeneous candidates for a job position. To capture such heterogeneity, we define the variable *ability* (or m), and make the following assumptions. The fraction of candidates in the society with abilities $m \in \mathbb{R}$ is given by the function $h(m)$. Assuming that the distribution of abilities is continuous, the fraction of candidates (or the probability that we select a candidate) with abilities equal to or lower than value d is given by:

$$\Pr[m \leq d] = H(d) = \int_{-\infty}^d h(m) dm$$

where $H(\infty) = 1$ and $H(-\infty) = 0$. It is common knowledge that this is the distribution of abilities in the society but not the particular abilities of each candidate.

Period 2. We assume that a firm opens a job position offering a contract to the candidates. The contract sets the selection intensity, an amount of training and the selection requirements.

Selection intensity refers to the information set that the firm will collect about the candidate's abilities (education background, work experience, ability or knowledge tests, interviews, etc.). Let us assume that all the information collected about the candidate can be summarized in one index, $I \in \mathbb{R}$. At that period, the firm knows only the distribution of abilities in the population and its relationship with I , described below. The fraction of available workers in the society (or the probability that we select a worker) with a value I is given by $g(I)$. Assuming again that the distribution of information is continuous, the percentage of people with a value I equal or lower to L is

$$\Pr[I \leq L] = G(L) = \int_{-\infty}^L g(I) dI = 1$$

where $G(\infty) = 1$ and $G(-\infty) = 0$.

To obtain a measure of selection intensity, we assume that the joint distribution of m and I is a bivariate normal distribution with $\mu_m = \mu_I = 0$, $\sigma_m = \sigma_I = 1$ and $\rho \geq 0$. In other words, we assume that $h(m)$ and $g(I)$ follow standardized normal probability distribution functions⁵ (pdf), $N(0, 1)$. It is well known (De Groot 2005, pp. 300) that the conditional distribution of abilities given that the information about the worker takes a value of I follows a normal distribution with an expected value $E(m/I) = \rho I$ and variance $V(m/I) = (1-\rho^2)$, $N(\rho I, 1-\rho^2)$, where ρ is the correlation coefficient between the information on the worker and the worker's abilities. We interpret ρ as the measure of selection intensity. Each firm has a known selection intensity that is exogenously determined. We consider only non-negative correlations, $0 \leq \rho \leq 1$ (note that when a variable X is negatively correlated, the simple transformation $I = -X$ will be positively correlated).

Period 3. Randomly, a candidate from the population of workers applies. Remember that although the abilities of a particular worker are unknown by the firm, their distribution is common knowledge. After the selection procedure, the firm obtains

⁵ Note that if the variables follow other normal distributions, these distributions can be easily standardized.

information on a particular candidate j (for example, the candidate has a high school degree with four years of experience, received 60 points in the ability or knowledge test and had a very good interview) that is summarized into one specific value identified as I_j .

Period 4. The firm decides to hire the candidate or not. When the candidate is not hired, the profits are zero.

Period 5. If the candidate is hired, the contract is executed. The firm provides the level of firm training t to the worker, and profits are generated. The firm profits $\pi(t)$ will be a function of the worker's abilities, m , training, t and a set of parameters (z, α, β) already defined in the literature review section:

$$\pi(t) = z\alpha t - \frac{zt^2}{2} + z\beta mt.$$

The firm is risk neutral and maximizes the expected profits generated by training, conditional on the information previously collected. For simplicity, we assume that the utility of workers is not affected by these workers' training levels. Then, the profits of the firms π are also the social surplus. In our model, we assume that firms capture the entire social surplus because we are not interested on the distribution of the social surplus, just in their maximization⁶.

Furthermore, it is difficult to assert that researchers can observe all the information firms collect about workers' abilities. The firms' information about a candidate includes, among other indicators, the level of education attained. However, researchers usually only have information about the level of education of the firm's workers as indicator of their abilities. To formally deduce the empirical implications of the model regarding the level of education, and without loss of generality, we assume that there are two levels of education – high (E) and low (e) – and all the workers with low education have values of I_j below⁷ L .

⁶ It is possible to change the workers' utility and firms' profit functions and establish negotiation rules that generate the same social surplus as this model does. Therefore, the results of the model are not driven by the fact that the workers' surplus is zero.

⁷ Then, $G(L)$ is the fraction of the population with low levels of education. Implicitly, we assume that the fraction of workers with low education when the value I is below L is 1, ($\Pr(e/I) = 1$), and zero when I is equal or higher than L , ($\Pr(e/I) = 0$). Alternatively, we can reproduce all our results assuming that the percentage of workers with low schooling does not increase with the workers' credentials; more formally,

The endogenous variable considered in the model is the level of firm-provided training t . We are not saying that the selection intensity and the level of education are chosen without considering that the firms will invest later in training. What we argue is that these decisions have been made before deciding on the training level of workers; thus, these decisions are exogenous at that moment, and we therefore do not model them.

B. Model solution and empirical implications

In period 4, the firm has to choose to hire the employee or not. Firms will hire the candidate if the expected profits are positive, conditional on the information previously collected:

$$E(\pi/I_j) = z\alpha t - zt^2/2 + z\beta E(m/I_j)t = z\alpha t - zt^2/2 + z\beta\rho I_j t > 0$$

Therefore, the candidate is rejected when $I_j \leq \frac{t-\alpha}{\beta\rho} = K$; otherwise, the candidate is accepted.

In period 2, the firm establishes the optimal level of training that maximizes the expected profits given the selection rule above:

$$\text{Max}_t: E(\pi) = (z\alpha t - zt^2/2 + z\beta E(m/I_j > K)t) (1 - \Pr(I_j \leq K))$$

The main results of the model are summarized in the following proposition (see Appendix A for details and proofs).

Proposition: Given a selection intensity ρ , we can establish the following functions and their properties:

- a) Function $K(\rho)$ is the minimum level required for a candidate to be selected. Given the optimal decisions, the minimum level of education required for hiring increases with selection intensity, $\frac{\partial K(\rho)}{\partial \rho} > 0$.

$\Pr(e/I)$ decreases or at least does not increase with I . However, this alternative way implies notation and algebra that complicates the exposition.

- b) Function $t(\rho)$ determines the average years of training of the workers selected by the firm. Given the optimal decisions, this function is increasing in the selection intensity, $\frac{\partial t(\rho)}{\partial \rho} > 0$.

In short, firms' selection intensity determines the firms' selection requirements and the investment in employees' training, given that all these variables are positively related. Regarding the selection requirements, we can observe the levels of education, which are expected to be positively related with the value of I . Remember that we have assumed that those workers with a value of I equal to or below L have low levels of education. Given Proposition a), we expect that when $K(\rho) \leq L$, all the workers who are not hired have low levels of education. When $K(\rho) > L$, all the workers with low levels of education are not hired by the firms.

The predictions of the model are summarized by the following hypotheses:

Hypothesis 1: The level of education is positively related to selection intensity ($\lambda_3 = \lambda > 0$).

Hypothesis 2: Without controlling for selection intensity (i.e., imposing $\delta_4 = 0$ in Equation (4)), the investments in training will be positively related to the level of education, ($\beta_4 = \beta_1 = \beta + \delta/\lambda = \delta/\lambda > 0$).

Hypothesis 3: The investments in training are positively related to selection intensity, ($\delta_3 = \delta > 0$).

Hypothesis 1 and Hypothesis 3 come directly from Propositions a) and b), respectively. Hypothesis 3 comes from the fact that the level of education is indirectly related to the training decisions because the fraction of workers with high levels of education hired by the firm increases when the firm has a higher selection intensity.

C. Discussion of the timing of the model

Implicitly, we have assumed that the training level is established in Period 2. Therefore, the training level is independent of the information about workers' abilities generated in Period 3, I_j . One could argue that firms can make better use of the information obtained in the selection process by making the training level contingent on

the information available about worker abilities (i.e. to make the decision after I_j is observed). In this case, the firm is going to maximize the following expected profits:

$$Max_t: E(\pi/I_j) = z\alpha t - \frac{zt^2}{2} + z\beta E(m/I_j)t - d = z\alpha t - \frac{zt^2}{2} + z\beta\rho I_j t - c$$

The parameter c reflects the (or difference in) costs of the contingent training contract. Take note that when $c = 0$, it is optimal for the firm to make use of the information available. However, other reasons could make it more attractive to establish the workers' training in Period 2 rather to implement a personalized training program; for example, it could be cheaper to implement because of all workers will be trained at the same time and by the same trainee, so we assume that ($c \geq 0$). Empirically, we cannot observe whether the firm implements a contingent training contract or not or the variability of training across the workers in the same firm. Therefore, our approach will be to compare the implications of each kind of training policy on the coefficients of Equations (3) and (4). From those estimations, we will make inferences about which training policy firms adopt most, or in other words, whether c is high or low.

From the first-order condition of the maximization problem above, one can deduce the following equation: $t^* = \alpha + \beta\rho I_j$. Given that workers with higher levels of education will have higher values of I_j on average, they will also receive more training than those workers with low levels of education when there are complementarities ($\beta > 0$) and there is information about workers' abilities ($\rho > 0$). Under the contingent training contract, we expect that the relationship postulated in Hypothesis 2 will be maintained after controlling for the selection intensity:

Hypothesis 4: After controlling for selection intensity, the investments in training are positively related to the level of education, ($\beta_4 = \beta > 0$).

In fact, these differences in training will increase with the information available or selection, ρ , so we expect that if we include in Equation (4) an interaction between schooling and selection intensity, the associated coefficient will be positive.

This alternative model remains silent on the relationship between education and selection intensity and our intuition is that Hypothesis 3 could be deduced by introducing

a certain degree of risk aversion in the agents, given that the risk introduced by contingent contracts is reduced when the selection intensity, ρ , increases.

Summing up, empirical support to Hypothesis 1 is considered as evidence that some firms use non-contingent training contracts, the described in Section B, implying that $\beta_4 < \beta_1$ but not excluding that other firms can use contingent training contracts ($\beta_4 > 0$). Whether firms use or not contingent training contracts is tested via Hypothesis 4.

III. Methods

A. Data description

The data for hypotheses testing are taken from a survey⁸ of Spanish industrial plants designed to obtain information on human resources and work organization practices. The original questionnaire was fine-tuned with a pre-test sample of 15 plant directors. The use of subjective assessments by the interviewee on various scales is a common practice in the empirical literature concerned with the analysis of human resources and work organization practices. This approach allows for the possibility to obtain information on certain concepts even if objective information is not available.

The target group was a collection of manufacturing plants in mainland⁹ Spain with 50 or more workers and whose economic activity was included in one of the 13 manufacturer sectors of the NACE¹⁰ classification for 1993. The unit of observation is the plant, not the firm as a whole. The sample of manufacturing plants was identified in CAMERDATA (the database for the chamber of commerce of Spain) and consisted of 3,000 plants. A stratified random sample that guaranteed the representativeness of strata by size and industrial sector based on 401 interviews (13.4% of the target group) was finally achieved. For each plant, a questionnaire form was completed between

⁸ The survey was jointly designed by a group of researchers from the Universitat Autònoma de Barcelona, Universitat Illes Balears, Universidad Pública de Navarra, and Universidad de Zaragoza.

⁹ The Canary and Balearic Islands and the two smallest Autonomous Communities in terms of GDP per capita, Castilla La Mancha and Extremadura, are not included in the sample.

¹⁰ The European Community statistical classification of economic activities.

December 2007 and April 2008 through personal interviews conducted by a specialized firm, in most cases, with the directors or with the production or human resources managers of the plant; each interview was approximately 60 minutes long.¹¹ Since some questionnaires were incomplete¹², we ended up with 362 observations. Table 1 compares the distribution of the plants by size and economic sector among the population of Spanish manufacturing plants¹³ and the sample. As seen, there are no important differences associated with size and economic sector.

—Insert Table 1 here—

B. Measures

Table 2 shows the distributions and the descriptive statistics of the variables used in the estimations. The original questions from which we define the variables are presented in Appendix B.

—Insert Table 2 here—

The dependent variables in Equations (3) and (4) are education attainment and investment in training, respectively. Regarding education, for each plant, we have information on the percentages of workers with no education (P_1 , 7.2% on average), with primary education (P_2 , 54.1% on average), with secondary education (P_3 , 27.2% on average %) and with university education (P_4 , 9.6% on average). We aggregate all this information in the variable $Education = P_1 + 2P_2 + 3P_3 + 4P_4$. This variable is the average of the standardized years of education of the workers of the plant when the differences between educational levels are held constant¹⁴. We do not have further information that

¹¹ Interviewer status was required by the questionnaire. Specifically, there are 9 possibilities: a single owner (1% of the sample), a partner or co-owner (3.4 % of the sample), a chairperson (2% of the sample), a director or general manager (13.8% of the sample), a sole director of a limited liability company (11.6% of the sample), a plant manager (9% of the sample), a production manager (13.5% of the sample), a human resources manager (17.8% of the sample) and others (32% of the sample).

¹² The missing values are spread throughout the different variables of the sample.

¹³ The population data are taken from the Central Directorate of Companies (Directorio Central de Empresas – DIRCE) of the Spanish National Institute of Statistics (Instituto Nacional de Estadística de España – INE).

¹⁴ Under this assumption, the number of years of education follows the equation: $Y = a + b y$, where a and b are positive parameters and y takes the value of 1 for those with no degree, 2 for those with primary education, 3 for those with secondary education and 4 for those with tertiary education. Therefore, the index y can be interpreted as a standardization of the years of education, $y = \frac{Y-a}{b}$, and the variable $Education$ can be interpreted as the average of this index y for the workers of the plant.

allows us to compute the real average years of education of the workers in the plant. For robustness, in Appendix C, we replicate the relevant part of the analyses by measuring education as the percentage of workers with more than a primary education (P4 + P3). The main conclusions are maintained. The other dependent variable, *Training*, is measured by ordered (five-point Likert scale) answers to the question related to investments in blue-collar workers' training.

The independent variable selection intensity, *Selection* from now on, is measured by ordered answers to the question related to the tools used during the selection process. To capture all the information available, four dummies must be included in the estimations. A potential problem with this approach is that some of the categories have very few observations, causing collinearity in the estimation. To avoid this problem and for the sake of expositional clarity, we transform the ordinal into a dummy variable that takes the value of 1 at or above the integer nearest the median of the ordinal variable and 0 otherwise (for a similar procedure, see Barrenechea-Mendez and Ben-Ner, 2017). Actually, with the data available we cannot assert that the workers trained today were hired using the current selection policies. Notwithstanding, this will be the case when policies such selection and training are quite stable along time.

We also use control variables such as *Job stability*, *Union influence*, *Technological level* and the introduction of *New products*, which are also measured by ordered (five-point Likert scale) answers. Thus, we transform these ordinal variables into dummy variables using the procedure described above. The recoded variables are shown in Column [3] of Table 2. Another set of control variables, *Collective agreements* and *Multinational*, are based on binary answers from the questionnaire, so they do not need to be transformed. Finally, the control variable *Size* is measured by the number of workers, so it is continuous. For the parameters associated with these control variables, we provide no prediction.

IV. Results

This section presents tests of the hypotheses about the estimations of Equations (3) and (4). Take note that given our data limitations, we do not provide evidence on the timing of the decisions and thus on the causality of the relationships tested.

Table 3 shows the results of the estimation of the *Education* determinants, Equation (3). Given the continuous nature of this dependent variable, we estimate the equation using OLS. The F-test rejects at the 1 percent level the null hypothesis that the parameter estimates of all the explanatory variables are zero. The estimate of *Selection* is positive and statistically significant at the 1 percent level. This result provides strong support for Hypothesis 1. Regarding the control variables, only *High technological level* and *New products* are associated with more *Education*.

—Insert Table 3 here—

The estimation of the *Training* equation, Equation (4), is relevant for Hypotheses 2, 3 and 4. Given the information available about training, we propose to estimate an ordered probit probability model. Initially, we estimate three different specifications. In Model 1, *Selection* is not included as an independent variable, while in Model 2, it is included. Model 3 also includes the interaction between *Selection* and *Education*.

Columns [1] to [3] in Table 4 shows the results. In the three alternative specifications, the likelihood ratio rejects at the 1 percent level the null hypothesis that the parameter estimates on all the explanatory variables are zero. In Model 1, the estimate on *Education* is positive and statistically significant at the 2 percent level. This result provides support for Hypothesis 2 and is consistent with the previous empirical work cited in the literature review.

In Models 2 and 3, the variable *Selection* is included. The estimates on this variable are positive and statistically significantly different from zero at the 1 percent level, providing support for Hypothesis 3. In these estimations, unlike the previous one, the estimate on *Education* is positive but not statistically significant at conventional levels of significance so Hypothesis 4 is rejected. The introduction of the variable *Selection* reduces the coefficient of *Education* (Model 1 vs 2) in two thirds, a reduction of two standard deviations which is statistically significant. In Model 3, we also include the interaction between *Selection* and *Education*, and the associated coefficient is not statistically significant.

Regarding the control variables, the statistically significant coefficients in the three models estimated are those for *Job stability*, *High technological level* and *Size*, which are positive in all cases.

—Insert Table 4 here—

A potential problem with the approach above is that it assumes that the error terms of Equations (3) and (4) are uncorrelated. If this is not the case, the estimates in Equation (4) will suffer from inconsistency and biasedness problems, i.e. endogeneity (Greene, 2012, pp. 265). There are methods for a simultaneous estimation of Equations (3) and (4) in which training is a continuous variable. Unfortunately, we are not aware of such methods when it is an ordered variable. Therefore, we transform the ordinal variable *Training* into a continuous one using the values of the Likert scale assigned by the interviewee. Before presenting the simultaneous estimations, we check that this transformation is innocuous. Columns [4] to [6] of Table 4 reproduce the estimations of Columns [1] to [3] using *Training* as a continuous variable. The results are very similar.

To test for the possible endogeneity of the variable *Education*, we estimate simultaneously Equations (3) and (4) using 3SLS (Zellner and Theil, 1962) and measuring *Training* as a continuous variable. To identify the model, we exclude a regressor from the *Education* equation, *New products*, the only variable (see Tables 3 and 4) correlated with *Education* but uncorrelated with *Training*. We present in Column [7] the results referred to Equation (4) while the results referred to Equation (3) are the same as the estimated in Table 3. Take note that in the absence of endogeneity, uncorrelated errors of the estimated equations, the OLS estimator will be more efficient than the 3SLS one (Greene 2012, pp. 372-3). The Hausman specification test cannot reject the null hypothesis of zero correlation between the error terms of Equations [3] and [4]. Therefore, Columns [1] to [3] are our preferred models. In any case, the main conclusions about the hypotheses are maintained.

V. Discussion and Conclusions

This paper presents evidence that Spanish industrial plants whose selection processes collect more information (a higher selection intensity) about blue-collar worker candidates, on average, select candidates with higher levels of education (an indicator of the selection requirements) and invest more in their training. Furthermore, after controlling by the differences in the plants' selection intensity the relationship between education and training is not statistically significant. In short, the evidence suggests that the plant's selection intensity is highly related to the plant's selection requirements (average education level of the selected workers) and provision of training.

This evidence is consistent with a model in which firms make explicit the requirements of the selection process jointly with the training associated with the job position before learning about the candidates for this position. At hiring, the firms consider whether the candidate will benefit from the training provided by the firm. Thus, it is expected that firms that offer higher levels of training will be more demanding or strict in the selection process, and vice versa. Consequently, the firm should establish a level of training or its equivalent level of requirements. Firms with more intense selection processes can discriminate better among the candidates, and therefore, they will be more demanding or exigent and provide more training to their workers.

The assumptions of the theoretical model are not specific to the environment analyzed, but obviously, further evidence in other environments and time periods is needed to corroborate the generality of the evidence presented. Furthermore, the model presented has other implications or assumptions not tested in the paper due to limitations of the data. For example, the timing of decisions, the type of contracts offered by the firms, differences in the workers' productivity or the rational and maximizing behavior of the agents. Further information, for example, details about the selection process and training of each worker of the firm or the productivity of the workers, could help to provide further tests of the arguments stated in the paper.

REFERENCES

Altonji, Joseph G. and James R. Spletzer. 1991. "Worker characteristics, job characteristics, and the receipt of on-the-job training." *Industrial and Labor Relations Review* 45 (1): 58-79.

Arulampalam, Wiji, Alison. L. Booth, and Peter Elias. 1996. "Work-related training and earnings growth for young men in Britain." Working Paper No. 440. University of Warwick, Department of Economics.

Arulampalam, Wiji, and Alison. L. Booth. 1998. "Training and labour market flexibility: is there a trade-off?" *British Journal of Industrial Relations* 36 (4): 521-536.

Barrenechea-Mendez, Marco A; and Avner Ben-Ner. 2017. "Mission Congruence and Organization Design: An Empirical Analysis of Childcare Facilities." *Industrial Relations: A Journal of Economy and Society* 56 (3): 411-426.

Barron, John M., Mark C. Berger, and Dan A. Black. 1997. "Employer search, training, and vacancy duration." *Economic Inquiry* 35 (1): 167-192.

Barron, John, M., Dan A. Black, and Mark A. Loewenstein. 1989. "Job matching and on-the-job training." *Journal of Labor Economics* 7 (1): 1-19.

Blundell, Richard, Lorraine Dearden, and Costas Meghir. 1996. The determinants and effects of work-related training in Britain (No. R50). IFS Reports, Institute for Fiscal Studies.

Booth, Alison. L. 1991. "Job-related formal training: who receives it and what is it worth?" *Oxford bulletin of economics and statistics* 53 (3): 281-294.

Booth, Alison. L. 1993. "Private sector training and graduate earnings." *The Review of Economics and Statistics* 75 (1): 164-170.

Brown, James N. 1989. "Why do wages increase with tenure? On-the-job training and life-cycle wage growth observed within firms." *The American Economic Review* 79 (5): 971-991.

De Groot, Morris H. 2005. *Optimal statistical decisions*. Vol. 82. John Wiley & Sons.

Gordon, Robert D. 1941. "Values of Mills' ratio of area to bounding ordinate and of the normal probability integral for large values of the argument." *The Annals of Mathematical Statistics* 12 (3): 364-366.

Goux, Dominique, and Eric Maurin. 2000. "Returns to firm-provided training: evidence from French worker-firm matched data." *Labour economics* 7 (1): 1-19.

Green, Francis. 1993. The determinants of training of male and female employees in Britain. *Oxford bulletin of economics and statistics* 55 (1): 103-122.

Greene, William H. 2012. *Econometric analysis*. Pearson Education India.

Greenhalgh, Christine, and George Mavrotas. 1994. "The role of career aspirations and financial constraints in individual access to vocational training." *Oxford economic papers* 46 (4): 579-604.

Greenhalgh, Christine, and Mark M. Stewart. 1987. "The effects and determinants of training." *Oxford Bulletin of Economics and Statistics* 49 (2): 171-190.

Groot, Wim, Joop Hartog, and Hessel Oosterbeek. 1994. Returns to within-company schooling of employees: the case of the Netherlands. In *Training and the private sector: international comparisons*, 299-388. University of Chicago Press.

Heckman, James. J. 2000. "Policies to foster human capital." *Research in economics* 54 (1): 3-56.

Kuckulenz, Anja, and Thomas Zwick. 2003. The impact of training on earnings: Differences between participant groups and training forms (No. 03-57). ZEW Discussion Papers.

- Lindley, Robert M.** 1975. "The demand for apprentice recruits by the engineering industry, 1951-71." *Scottish Journal of Political Economy* 22 (1): 1-24.
- Lillard, Lee A., and Hong W. Tan.** 2012. Private sector training: who gets it and what are its effects? In 35th Anniversary Retrospective, Vol. 35, 247-308. Emerald Group Publishing Limited.
- Lynch, Lisa M.** 1992. "Private-sector training and the earnings of young workers." *The American Economic Review* 82 (1): 299-312.
- Lynch, Lisa M., and Sandra E. Black.** 1998. Beyond the incidence of employer-provided training. *Industrial and Labor Relations Review* 52 (1): 64-81.
- Mason, Geoff, Penny Mok, Peter Nunns, Philip Stevens, and Jason Timmins.** 2012. To Make or Buy (Skills): An analysis of training decisions using microdata.
- Mathis, Robert L., John H. Jackson and Sean R. Valentine,** Human Resource Management, 14th Edition. 2014. Cengage Learning.
- Merrilees, William J.** 1980. Alternative models of apprentice recruitment: with special reference to the British engineering industry. *Applied Economics* 15 (1): 1-21.
- Mincer, Jacob.** 1994. Human capital: A review. Labor economics and industrial relations: Markets and institutions, 109-141.
- Oosterbeek, Hessel.** 1996. "A decomposition of training probabilities." *Applied Economics* 28 (7): 799-805.
- Oosterbeek, Hessel.** 1998. "Unravelling supply and demand factors in work-related training." *Oxford economic papers* 50 (2): 266-283.
- Pannenberg, Markus.** 1996. Financing on-the-job training: Shared investment or promotion based system? Evidence from Germany (No. 135). DIW Discussion Papers.
- Pischke, Jörn-Steffen.** 2001. "Continuous training in Germany." *Journal of population economics* 14 (3): 523-548.

Royalty, Anne Beeson. 1996. "The effects of job turnover on the training of men and women." *Industrial and Labor Relations Review* 49 (3): 506-521.

Rosen, Sherwin. 1976. "A theory of life earnings." *Journal of Political Economy* 84 (4) (Part 2): S45-S67.

Stevens, Margaret. 1994. "A theoretical model of on-the-job training with imperfect competition." *Oxford economic papers* 46 (4): 537-62.

Tan, Hong. W., Bruce Chapman., Christine E. Peterson, and Alison Booth. 1991. Youth Training in the United States, Britain, and Australia, in *Research in Labor Economics*, Vol. 13, 63-99. JAI Press.

Zellner, A., & Theil, H. 1962. Three-stage least squares: simultaneous estimation of simultaneous equations. *Econometrica: Journal of the Econometric Society*, 54-78.

Tables

Table 1—Distribution of the Plants by Size and Economic Sector

Variable	Category	% Sample	% Population
<i>Size</i>	From 50 to 99 employees	48.62	55.07
	From 100 to 199 employees	32.60	24.39
	From 200 to 499 employees	13.81	15.22
	More than 500 employees	4.97	5.32
	Total	100	100
<i>Industry</i>	Food, drink and tobacco	17.12	16.05
	Textile industry, dressmaking, leather and footwear	8.01	6.39
	Wood, cork, paper and graphic arts	10.22	10.96
	Furniture and various manufacturing industries	4.14	5.36
	Rubber, plastic materials and non-metallic mineral products	19.61	16.21
	Metallurgy metal equipment (excluded machinery)	12.43	16.98
	Chemical industry	3.04	5.54
	Mechanical equipment and machinery	8.56	8.71
	Electric equipment	3.04	3.86
	Motor vehicle and transport supply	8.84	4.63
	Electronic, medical, optical and computer equipment	2.49	2.78
	Pharmaceutical industry	1.93	2.01
	Aeronautical industry	0.55	0.48
	Total	100	100

Notes. The Spanish plant population is taken from the Central Directorate of Companies. In the sample, the variable size is continuous but in the population, the information about the size appears in categories. To make them comparable, we split the sample into four groups.

Table 2—Frequency Distribution and Descriptive Statistics

Variable [1]	Original codification [2]	Used codification [3]	Mean [4]	S.d. [5]	
<i>Training</i>	<i>Likert scale</i>	<i>Original one</i>			
	<i>Nil or Very Low</i>		0.30		
	<i>Low</i>		0.12		
	<i>Average</i>		0.56		
	<i>High</i>		0.25		
<i>Education</i>	<i>Very High</i>	0.33			
	<i>Percentages</i>	<i>Continuous</i> <i>Standardized years</i>			
	<i>% No studies</i>		0.07 (0.16)	2.40	0.43
	<i>% Primary</i>		0.55 (0.27)		
	<i>% High school</i>		0.28 (0.24)		
<i>% University</i>	0.10 (0.13)				
<i>Selection</i>	<i>Likert scale</i>	<i>Dummy</i> <i>High selection</i>			
	<i>Nil or very low</i>		0.05	0.79	
	<i>Low</i>		0.16		
	<i>Average</i>		0.62		
	<i>High</i>		0.17		
<i>Job stability</i>	<i>Very High</i>	0.00			
	<i>Likert scale</i>	<i>Dummy</i> <i>High job stability</i>			
	<i>Nil or very low</i>		0.02	0.50	
	<i>Low</i>		0.04		
	<i>Average</i>		0.44		
<i>High</i>	0.43				
<i>Unions</i>	<i>Very High</i>	0.08			
	<i>Likert scale</i>	<i>Dummy</i> <i>High influence</i>			
	<i>Nil or very low</i>		0.06	0.67	
	<i>Low</i>		0.18		
	<i>Average</i>		0.44		
<i>High</i>	0.25				
<i>Collective agreements</i>	<i>Very High</i>	0.07			
	<i>Dummy</i>	<i>Original one</i> <i>Original one</i> <i>Dummy</i> <i>High Technology</i>	0.57		
	<i>Multinational</i>		0.29		
	<i>Technological level</i>				
	<i>Low</i>		0.40	0.28	
<i>Middle low</i>	0.32				
<i>Size</i>	<i>Middle high</i>	0.23			
	<i>High</i>	0.05			
	<i>Continuous</i>	<i>Original one</i> <i>Dummy</i> <i>High</i>	201.21	523.71	
	<i>New products</i>				
	<i>Likert scale</i>				
<i>Nil or very low</i>	0.09		0.57		
<i>Low</i>	0.33				
<i>Very High</i>	<i>Average</i>	0.20			
	<i>High</i>	0.30			
	<i>Very High</i>	0.07			

Notes. The number of observations is 362. Column [1] displays the variables used in the estimations. Column [2] presents the percentages of each category (and the standard deviations in parentheses). Column [3] presents the recodification of the original variables according to the criterion explained in the text, while columns [4] and [5] present their respective descriptive statistics. The unit of observation is the plant.

Table 3—Determinants of Education

Independent variables	OLS
<i>Selection</i>	25.70 (4.90)
<i>Job stability</i>	-0.17 (4.29)
<i>Unions</i>	-0.08 (4.40)
<i>Collective agreements</i>	-2.66 (4.31)
<i>Multinational</i>	2.58 (4.79)
<i>High technological level</i>	23.46 (4.93)
<i>Size</i>	0.002 (0.002)
<i>New products</i>	8.64 (4.58)
Constant	213.77 (11.58)
<i>R</i> ²	0.15
<i>F</i>	10.21
<i>Observations</i>	362

Notes: Ordinary least squares (OLS) estimates (with robust standard errors). Standard deviations are in parentheses.

Table 4—Determinants of Training

Independent variables	Ordered probit			OLS			3SLS
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 2
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
<i>Selection</i>		0.96 (0.18)	2.18 (0.84)		0.62 (0.11)	1.40 (0.55)	0.58 (0.25)
<i>Education</i>	0.003 (0.001)	0.001 (0.001)	0.006 (0.003)	0.002 (.001)	0.001 (0.001)	0.004 (0.002)	0.002 (0.01)
<i>Selection* Education</i>			-0.006 (0.004)			-0.004 (0.002)	
<i>Job stability</i>	0.46 (0.12)	0.47 (0.12)	0.49 (0.12)	0.30 (0.08)	0.29 (0.07)	0.30 (0.07)	0.29 (0.07)
<i>Unions</i>	-0.05 (0.13)	-0.05 (0.13)	-0.06 (0.13)	-0.03 (0.08)	-0.03 (0.08)	-0.03 (0.08)	-0.03 (0.8)
<i>Collective agreements</i>	0.14 (0.12)	0.10 (0.12)	0.10 (0.12)	0.10 (0.08)	0.08 (0.08)	0.07 (0.08)	0.08 (0.08)
<i>Multinational</i>	0.02 (0.12)	-0.01 (0.12)	0.004 (0.12)	0.03 (0.08)	0.01 (0.08)	0.01 (0.08)	-0.01 (0.09)
<i>High technological level</i>	0.24 (0.14)	0.25 (0.14)	0.25 (0.14)	0.16 (0.09)	0.16 (0.8)	0.16 (0.08)	0.12 (0.22)
<i>Size</i>	0.001 (0.0002)	0.001 (0.0002)	0.001 (0.0002)	0.0003 (0.0001)	0.0003 (0.0001)	0.0003 (0.0001)	0.0003 (0.0001)
<i>New products</i>	0.07 (0.12)	0.001 (0.13)	-0.006 (0.13)	0.07 (0.08)	0.01 (0.8)	0.01 (0.8)	
<i>Constant</i>				2.34 (0.27)	2.22 (0.25)	1.59 (0.49)	1.88 (1.92)
<i>Cut1</i>	-0.71	-0.67	0.30				
<i>Cut2</i>	0.16	0.33	1.31				
<i>Cut3</i>	1.88	2.17	3.15				
<i>Cut4</i>	3.36	3.68	4.67				
<i>Pseudo R2</i>	0.07	0.11	0.12				
<i>Log likelihood</i>	-385.83	-366.91	-365.87				
<i>R2</i>				0.13	0.23	0.23	0.22
<i>F</i>				7.81	11.11	10.33	
<i>Wald X2</i>	52.96	79.30	83.01				
<i>X2</i>							104.88
<i>N</i>	362	362	362	362	362	362	362

Notes. Columns [1] to [3] report Ordered probit (robust standard errors) estimates. Columns [4] to [6] report OLS (robust standard errors) estimates. Column [7] reports a single-equation instrumental-variables regression (2SLS) (robust standard error) estimates. Standard deviations are in parentheses.

Appendix A: The demonstration of the proposition

The problem of the firm can be written as

$$\text{Max}_t: E(\pi) = (z\alpha t - zt^2/2 + z\beta E(m/I_j > K)t) (1 - \Pr(I_j \leq K))$$

Given that

$G(K) = \Pr(I_i \leq K) = \Phi(K)$, where $\Phi(\cdot)$ the cumulative distribution function (cdf) of a random variable with a normal distribution function (0, 1).

$$E(m/I_j > K) = \int_K^\infty \rho I g(I) dI = \rho \int_K^\infty I g(I) dI = \rho \lambda(K),$$

$\int_K^\infty I g(I) dI$ is the mean of a truncated normal distribution $N(0,1)$. $\lambda(K) = \frac{\phi(K)}{(1-\Phi(K))} = \frac{1}{R(K)}$ is the inverse of the Mills ratio $R(K)$ and $g(K) = \phi(K)$, where $\phi(\cdot)$ is the probability distribution function (pdf) of a normal distribution with a mean of 0 and a standard deviation of 1. We use the following properties of these functions, $\frac{\phi'(K)}{\phi(K)} = -K$, $\lim_{K \rightarrow \infty} \lambda(K) = K$, $0 < \lambda(K)(\lambda(K) - K) < 1$. For further details, see Gordon (1941).

Additionally, taking into account that $K = \frac{t-\alpha}{\beta\rho}$, the problem above can be reformulated in the following way:

$$\text{Max}_K: E(\pi) = 2z\beta\rho(K\beta\rho + \alpha)(\lambda(K) - K) (1 - \Phi(K)) =$$

$$2z\beta\rho(K\beta\rho + \alpha)(\phi(K) - K (1 - \Phi(K)))$$

Note that when $t=0$, $E(\pi) = 0$, therefore $K > -\frac{\alpha}{\rho\beta}$.

The first- and second-order conditions for the existence of an internal optimum are, respectively,

$$\frac{dE(\pi)}{dK} = 0$$

$$\frac{d^2E(\pi)}{dK^2} < 0$$

The envelope theorem establishes the changes in the optimum K^* due to changes in the intensity selection:

$$\frac{dK^*}{d\rho} = -\frac{\frac{dE(\pi)}{dKd\rho}}{\frac{d^2E(\pi)}{dK^2}}$$

Proposition a: If there is an optimum, then $\frac{dK^*}{d\rho} > 0$.

Proofs: From the second-order condition: $\frac{d^2E(\pi)}{dK^2} < 0$. Then, accordingly with the envelope theorem, $\frac{dE(\pi)}{dKd\rho} > 0$ must be demonstrated.

$$\frac{dE(\pi)}{dK} = 2z\beta\rho((\beta(\phi(K) - K(1 - \Phi(K))) + (\rho\beta K + \alpha)(\phi'(K) - (1 - \Phi(K)) + K\phi(K))).$$

Taking into account that $\frac{\phi'(K)}{\phi(K)} = -K$, and using some algebra, we can deduce

$$\frac{dE(\pi)}{dK} = 2z\beta\rho\phi(K) (\rho\beta(1 - 2KR(K)) - \alpha(R(K)))$$

Deriving by ρ ,

$$\begin{aligned} \frac{dE(\pi)}{dKd\rho} &= 2z\beta\phi(K) (\rho\beta(1 - 2KR(K)) - \alpha(R(K))) \\ &\quad + 2z\beta\rho\phi(K) (\beta(1 - 2KR(K))) \end{aligned}$$

The first-order condition will be fulfilled when

$$\rho\beta(1 - 2KR(K)) - \alpha(R(K)) = 0$$

Given that $\alpha(R(K)) > 0$, then $(1 - 2KR(K)) > 0$ must occur.

Therefore, in the optimum: $\frac{dE(\pi)}{dKd\rho} = 2z\beta\rho\phi(K) (\beta(1 - 2KR(K))) > 0$. QED

Proposition b: If there is an optimum, then $\frac{dt^*}{d\rho} > 0$.

Proofs: By definition, $K = \frac{t - \alpha}{\beta\rho}$, so $t^* = 2(\rho\beta K^* + \alpha)$.

Therefore, $\frac{dt^*}{d\rho} = 2(\beta K^* + \rho\beta \frac{dK^*}{d\rho}) > 0$. QED

This is a unique interior solution.

From the first-order condition $0 = \rho\beta(1 - 2KR(K)) - \alpha(R(K))$, one can establish that $\frac{\alpha}{\rho\beta} = \frac{1}{R(K)} - 2K = M(K)$

It is easy to show that $M(-\frac{\alpha}{\rho\beta}) = \frac{1}{R(-\frac{\alpha}{\rho\beta})} + 2\frac{\alpha}{\rho\beta} > \frac{\alpha}{\rho\beta}$ and $\lim_{K \rightarrow \infty} M(K) = -K = -\infty$, as $\lim_{K \rightarrow \infty} 1/R(K) = K$. Given that $0 < \lambda(K)(\lambda(K) - K) = \frac{1-KR(K)}{(R(K))^2} < 1$, $\frac{dM(K)}{dK} = \frac{1-KR(K)}{(R(K))^2} - 2 < -1 < 0$ and guarantees the existence of a unique interior solution $K^* > -\frac{\alpha}{\rho\beta}$. Therefore, $K^* = K(\rho) = M^{-1}(\frac{\alpha}{\rho\beta})$. QED

The solution is a maximum:

$$\begin{aligned} \frac{d^2 E(\pi)}{dK^2} = & 2z\beta\rho\phi'(K) (\rho\beta(1 - 2KR(K)) - \alpha(R(K))) \\ & + 2z\beta\rho\phi(K) (\rho\beta(-2R(K) - 2K(KR(K) - 1)) - \alpha(KR(K) - 1)) \end{aligned}$$

The second-order condition implies that this is a maximum when

$$\rho\beta(-2R(K) - 2K(KR(K) - 1)) - \alpha(KR(K) - 1) < 0$$

Multiplying this expression by $1/\rho\beta R(K)$ and replacing $\frac{\alpha}{\rho\beta} = \frac{1}{R(K)} - 2K$, the inequality above is equivalent to: $\frac{1-KR(K)}{(R(K))^2} - 2 < -1 < 0$. QED

Appendix B: Survey items used for the variables

Training

Please indicate the degree of suitability of the following human resources management practices among your workers (1: Nil or very low, 5: Very high)

	Nil or very low	Low	Average	High	Very high
The investment in blue-collar workers' training, in hours and in money, is...	1	2	3	4	5

Education

Please indicate the percentage of blue-collar workers in your plant that have the education listed.

%
No studies
Primary education
High school
University

Selection

Please indicate the degree of application of the following human resources management practices among your blue-collar workers (1: nil or very low; 5 very high).

	Nil or very low	Low	Average	High	Very high
The variety of tools used during the selection process (interviews, personality and ability tests, simulations) is ...	1	2	3	4	5

Job stability

Please indicate the degree of application of the following human resources practices among your workers (1: Nil or very low, 5: Very high).

	Nil or very low	Low	Normal	High	Very high
The commitment to indefinitely keep the employment relationship with our workers is...	1	2	3	4	5

Unions

How do you assess the influence of unions on your firm's blue-collar workers?

1. Very low, 2. Low, 3. Average, 4. High, 5. Very High.

Collective agreements

Does there exist a plant- or firm-specific collective agreement that regulates the working conditions of your production workers?

1. Yes, 0. No.

Multinational

Does the parent company have any other production plant in foreign countries (outside of Spain)?

1. Yes, 0. No.

Technological level

Economic sector	Technological intensity
Aeronautical industry	
Pharmaceutical industry	
Electronic, medical, optical and computer equipment	High
Chemical industry	
Machinery and metal equipment	
Electric equipment	Middle high
Motor vehicle and transport supply	
Rubber, plastic materials and non-metallic mineral products	Middle low
Metallurgy	
Food, drink and tobacco	

Textile industry, dressmaking, leather and footwear

Low

Wood, cork, paper and graphic arts

Furniture and various manufacturing industries

Size

Approximately how many workers did you have in 2005?

Instrument (new products)

Please indicate your degree of agreement with the following assertions (From 1: Total disagreement to 5: Total agreement)

	Total disagreement	Disagreement	Neither agreement nor disagreement	Agreement	Total agreement
Many new products are introduced every year	1	2	3	4	5

Appendix C: Robustness Analysis

For robustness, we reproduce Table 3 (Table 3A) and the three first columns of Table 4 (Table 4A) using an alternative measure of education: the percentage of workers with a level of education higher than primary education.

The main difference is related with-to Hypothesis 4. The significance of the coefficient associated with the variable *Education* is strongly reduced when we control for *Selection* (Model 2), but its significance is now around 10%. In Model 3, the interaction is negative and statistically significant at 5%. In Table 4A, the control variable *High technological level* is not statistically significant, with a p-value of 13%.

Table 3A—Determinants of Education

Independent variables	OLS
<i>Selection</i>	15.23 (2.91)
<i>Job stability</i>	-0.06 (2.68)
<i>Unions</i>	0.69 (2.87)
<i>Collective agreements</i>	1.69 (2.72)
<i>Multinational</i>	1.00 (3.07)
<i>High technological level</i>	16.56 (3.35)
<i>Size</i>	0.004 (0.002)
<i>New products</i>	1.85 (2.86)
<i>Constant</i>	19.58 (7.01)
<i>R²</i>	0.16
<i>F</i>	10.82
<i>Observations</i>	362

Notes: OLS (robust standard error) estimates using an alternative measure of education. Standard deviations are in parentheses.

Table 4A—Determinants of Training

Independent variables	Ordered probit		
	Model 1	Model 2	Model 3
	[1]	[2]	[3]
<i>Selection</i>		0.93 (0.18)	1.12 (0.25)
<i>Education</i>	0.007 (0.002)	0.004 (0.002)	0.01 (0.01)
<i>Selection*Education</i>			-0.01 (0.01)
<i>Job stability</i>	0.46 (0.12)	0.47 (0.12)	0.49 (0.12)
<i>Unions</i>	-0.07 (0.13)	-0.06 (0.13)	-0.06 (0.13)
<i>Collective agreements</i>	0.11 (0.12)	0.09 (0.12)	0.10 (0.12)
<i>Multinational</i>	0.03 (0.12)	-0.01 (0.12)	0.01 (0.12)
<i>High technological level</i>	0.21 (0.13)	0.21 (0.13)	0.21 (0.13)
<i>Size</i>	0.001 (0.002)	0.001 (0.0002)	0.001 (0.0002)
<i>Cut1</i>	-1.54	-1.02	-0.69
<i>Cut2</i>	-0.67	-0.03	0.31
<i>Cut3</i>	1.05	1.82	2.16
<i>Cut4</i>	2.55	3.34	3.68
<i>Pseudo R2</i>	0.07	0.12	0.12
<i>Wald X2</i>	57.80	82.59	83.22
<i>Log likelihood</i>	-384.18	-365.81	-365.18
<i>N</i>	362	362	362

Notes: Ordered probit (robust standard errors) estimates using an alternative measure of education. Standard deviations are in parentheses.