Barrenechea-Méndez, Marco; Ortín Ángel, Pedro; Rodes-Mayor, Eduardo C.

Article

Employee selection, education, and firm-provided training

BRQ Business Research Quarterly

Provided in Cooperation with:
Asociación Científica de Economía y Dirección de Empresas (ACEDE), Madrid

Suggested Citation: Barrenechea-Méndez, Marco; Ortín Ángel, Pedro; Rodes-Mayor, Eduardo C. (2022) : Employee selection, education, and firm-provided training, BRQ Business Research Quarterly, ISSN 2340-9436, Sage Publishing, London, Vol. 25, Iss. 3, pp. 224-242, https://doi.org/10.1177/2340944420970960

This Version is available at:
https://hdl.handle.net/10419/261925

Regular Paper

Business Research Quarterly
2022, Vol. 25(3) 224–242

DOI: 10.1177/2340944420970960
journals.sagepub.com/home/brq
SAGE

Employee selection, education, and firm-provided training

Marco Barrenechea-Méndez[1], Pedro Ortín-Ángel[2] and Eduardo C. Rodes[2]

Abstract

Theoretical debate suggests at least three strategies for firms to provide training to employees in the same job position: individualized and egalitarian with or without adaptation to the abilities of the recruited employees. The article provides a formal framework for deriving distinctive empirical implications regarding the relationship of these strategies with the firms' selection policies, which are tested using a dataset of blue-collar workers in Spanish industrial plants. The evidence is consistent with the empirical implications of the egalitarian strategy with adaptation. This strategy entails providing the same level of training to all workers in the same job position and setting this level according to the average ability of recruited workers. Paradoxically, this strategy has not been used to interpret the results of the existing empirical literature.

JEL CLASSIFICATION: M53; M54; M21; J24

Keywords

Training, selection and education

Introduction

Barron et al. (1989, 1997) pointed out a positive relationship between firm-provided training and the intensity of their workers' selection procedures. We are not aware of any posterior academic effort to reinvestigate that relationship until now. By contrast, there is extensive literature analyzing other determinants of firm-provided training, predominantly, the workers' educational level (see Kramer & Tamm, 2018). Although we provide evidence of a positive relationship between firm-provided training and workers' education level, this relationship vanishes when we control for the firms' selection intensity. This is a result not predicted by the previous literature. Furthermore, in our empirical application selection intensity emerges as the most important variable to explain training. This evidence suggests the need to revisit the study of the relationship between training and selection intensity. This article is an effort to analyze the relationships among firm-provided training, the selection process intensity, and workers' education in light of theoretical advances that have occurred in recent years.

Barron et al. (1997) developed a formal model analyzing the relationship between the training provided by firms and their selection procedures. In this model, training is an exogenous variable; therefore, these authors are implicitly assuming that the firms' training strategy is (1) egalitarian: firms offer the same level of training to all workers in the same job position, thus they do not use the information provided by the selection process for training decisions and (2) static: firms do not adapt the training level to the relevant information about the workers' abilities that might be revealed between the moment in which the candidate is hired and training is provided.

These assumptions contrast with those usually made in models interpreting the empirical evidence on the determinants of firm-provided training (see Kramer & Tamm, 2018, for a recent contribution). The most common finding of this literature has been the positive relationship between workers' educational level (which is interpreted as an indicator of

[1]Public University of Navarre, Pamplona, Spain
[2]Autonomous University of Barcelona, Barcelona, Spain

Corresponding author:
Marco Barrenechea-Méndez, Public University of Navarre, 31006 Pamplona, Spain.
Email: marco.barrenechea@unavarra.es

workers' abilities) and the training provided by firms. This evidence has been interpreted as indicative of the existence of dynamic complementarities between training and ability (Cunha & Heckman, 2007), which implies that firms find it profitable to set the level of training according to the ability of the candidates (Cunha & Heckman, 2007; Heckman, 2000; Mincer, 1992; Rosen, 1976). Implicitly, this literature is assuming that the firms' training strategy is (1) individualized, as opposed to egalitarian, and (2) adaptative, that is, firms adapt the training level to the information about the workers' abilities. Following Prasad and Tran (2013), we will term direct contract to a strategy combining these two characteristics.

Comparing these two strategies, we observe that the second one, direct contract, is less restrictive (individualized and adaptative vs egalitarian and static) and, one would expect, more efficient (see Bishop, 1991, for a discussion on this argument). Therefore, further theoretical justifications need to be provided to explain why firms may want to self-restrict their training strategies. Henceforth, we refer to the training strategy with egalitarian and static characteristics as the restrictive contract. Psychological literatures related to procedural justice (Folger & Cropanzano, 1998; Greenberg, 1996; Wiesenfeld et al., 2007) and social comparison (Ployhart et al., 2006; Taylor et al., 1990; Wood, 1989) provide informal arguments for why human resources practices cannot be individualized.

In addition, contractual justifications have been provided recently in the literature on the economics of private-sector training. Since the influential contributions of Acemoglu and Pischke (1998, 1999a, 1999b) explained why firms might provide general training to their workers, the focus of the theoretical debate has moved from the distinction between general and specific training toward the implications of labor market imperfections. For example, Leuven's (2005) literature review classifies the theoretical models related to the economics of private-sector training based on their assumptions about competition in the labor market, or, in other words, whether it is perfect or imperfect (Boom, 2005).

The main source of imperfections in labor markets considered in the theoretical literature is the information about the workers' abilities. Several authors (Chang & Wang, 1995; Kahn & Hubberman, 1988; Prasad & Tran, 2013; Prendergast, 1993) argue that it is generally not feasible to make training contracts contingent on workers' abilities and that firms should therefore commit to providing the same level of training to all workers with the same contractible characteristics, such as those in the same job position. In fact, Prasad and Tran (2013) stress the importance of adapting training policies to environmental changes. Based on this literature, we suggest a third training strategy, one that is (1) egalitarian and that (2) adapts the level of training as the firm updates its information about the workers' abilities gathered after the firm's selection process. We term this training strategy an indirect contract.

This article contributes to the literature in the following ways. First, it provides a formal analysis of the implications of indirect training contracts on selection procedures. Second, it develops a theoretical framework encompassing and comparing the empirical implications of the three firm training strategies or contracts described above. Finally, it provides evidence on these implications, using cross-sectional data based on a survey of Spanish industrial plants providing information on human resources policies for blue-collar workers. The results are consistent with the implications of indirect contracts rather than with those of the other two types of contracts, direct and restrictive.

The remainder of this article is organized as follows. Section "Theoretical framework" provides the analysis of the empirical implications of the three training strategies or contracts already mentioned. Section "Methods" details the methods for providing the evidence. Section "Results" presents the results and section "Discussion and managerial implications" concludes the article and discusses its managerial implications.

Theoretical framework

Scope of the framework

This section develops a general framework to analyze the relationship between selection policies and firm-provided training under the three training strategies or contracts mentioned in the introduction. Therefore, the general assumptions of this framework are common to these three contracts and reflect the existence of imperfect information regarding candidates' abilities as well as the complementarities between those abilities and firm-provided training. The timing of the model reproduces the sequence of actions made by firms. First, firms evaluate available candidates using different levels of intensity (including null intensity). Then, some of those candidates are selected and join the firms. Finally, firms provide training to those employees.

The differences among the contracts (or firm training strategies) are discussed in section "Types of contracts," and this discussion is mainly focused on deriving (section "Empirical implications and the role of education") the implications that can be tested using the data available. Obviously, the framework is a simplification of the related literature and some important aspects of their discussions are therefore missing. Furthermore, we focus on two properties of the training strategy (individualization and adaptation), which in most cases are not central in this literature.

A general framework

Period 1: abilities. A pool of heterogeneous candidates for a job position reaches a firm. To capture this heterogeneity, we define the variable *ability* $(m \in \mathbb{R})$. Note that m does

not represent the candidates' innate abilities but the candidates' abilities at the time that the job position is open. We do not analyze how those abilities are produced or the possible components of those abilities (e.g., innate or acquired, manual or intellectual, specific or general). We assume that there exist differences in abilities and that agents have an idea of how such abilities are distributed in the population (e.g., due to their previous experience). To capture that, we assume that the workers' abilities in society follow a standardized normal probability distribution function, $h(m)$, which is common knowledge. Note, however, that although the agents know the distribution of abilities, initially, they do not know the ability of a particular candidate.

Period 2: information collection. The firm opens a job position with an associated selection procedure. Each job position opened is understood as an evaluation process of a particular candidate. Following Barron et al. (1997), we focus on two characteristics of the selection procedure, its intensity and extensity. Intensity is the set of tools for acquiring information about the candidate's ability (e.g., credentials, ability or knowledge tests, interviews, Internet searches) that determine the quality of the information collected. Based on this information, the firm decides whether to reject or hire the candidate. Extensity is the number of times that the process is repeated until a candidate is hired and, consequently, it is directly related to the probability of rejecting a candidate. For the sake of simplicity, in this period, we focus on intensity and defer the analysis of extensity to Period 3. In both cases, the goal is the same: to set the assumptions to represent these concepts in terms of a single variable. Next, we detail the assumptions related to selection intensity.

First, the selection procedure generates information about the candidates' abilities. After this process, the firm will have more (or at least the same) information about the candidate's abilities than before the process. This information is summarized in the index number $I \epsilon \mathbb{R}$, which also follows a standardized normal probability distribution function $g(I)$. Second, the information about the candidates' abilities I could have different qualities. To obtain a measure of the quality of the information, we benefit from the work of De Groot (2005) on normal distributions. 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 $\rho \geq 0$ (then $h(m)$ and $g(I)$ are the marginal distribution functions). Therefore, the conditional distribution of abilities given the information available about the worker I, follows a normal distribution with an expected value $E(m / I) = \rho I$ and variance $V(m / I) = (1 - \rho^2)$, where ρ is the correlation coefficient between the information collected I and the worker's true abilities m. Therefore, we interpret ρ as the measure of the information quality. We consider only nonnegative correlations,[1] $0 \leqslant \rho \leqslant 1$.

Third, information quality ρ can be improved by investing more into the selection process (using more or better selection tools) and, consequently, increasing the costs of the firm. Then, we define function $C(\rho)$, which represents the minimum cost at which the firm can obtain quality ρ. We assume that $C(\rho)$ is increasing in ρ and that it is always possible to contract with a candidate without a selection process, $C(0) = 0$. In addition, we assume that $C(\rho)$ is the same for all the firms, which can be justified by the fact that they usually use similar tools for collecting information about the candidate's ability. From now on, we refer to ρ as selection intensity.

Finally, selection intensity affects the social surplus associated with the job position. For the sake of simplicity, we assume that the agents (firms and candidates) are risk neutral and maximize the expected social surplus—the difference between the expected profits of the job position p_i^c and its costs: $S_i^c = p_i^c - C(\rho)$, where subindex i refers to a particular firm and superscript c to the type of the contract: direct, indirect, or restrictive. The profits associated with the job position depend on exogenous and endogenous variables introduced (and described jointly with p_i^c) in the next periods and in sections "Types of contracts" and "Empirical implications and the role of education." In our analysis, the surplus-sharing process does not play any role; hence, we do not model it.[2]

In sum, at Period 2, a random candidate from the pool applies to a job position opened by a firm and goes through the selection process. Every selection process has an associated selection intensity ρ. At this point, we emphasize that the selection intensity can be never established or modified after Period 2. Then, it could be the cause, but never consequence of the decisions made after Period 2.

Period 3: candidate selection. Now we focus on the extensity of the selection procedure or selection criteria. We assume that each firm establishes L, the minimum expected ability of a candidate for being accepted by firm i. At the beginning of this period, firm i obtains information on the particular candidate j (for example, the candidate has a high school degree with 4 years of experience, scored 60 points on the ability or knowledge test, and had a very good interview) that is summarized into I_j. The candidate's expected ability is $\widehat{m_j} = E(m / I_j) = \rho I_j$, so when $I_j > L / \rho$ the candidate is hired. In other words, the firm establishes the probability of rejecting a given candidate from a set of candidates that apply randomly to the job position $\Phi(L / \rho)$, where $\Phi(.)$ represents the cumulative distribution function of a standard normal distribution. For example, if the minimum expected ability is below the average (L is negative) and $\rho = 0$, then $\Phi(-\infty) = 0$ and nobody is rejected. Since the probability that a candidate is rejected strictly increases with $K = L / \rho$, we interpret K as the extensity of the selection process.

We assume that a rejected candidate will generate profits of zero, whereas a candidate filling the job position will generate expected profits of $E(\pi_i)$. Therefore, before obtaining any information about the candidate's abilities, the expected profits associated with the job position p_i^c will be profits $E(\pi_i)$ times the probability of filling the job position $(1-\Phi(L/\rho)): p_i^c = E(\pi_i)(1-\Phi(L/\rho))$.

Although the profits of filling the job position π_i will be defined in the next period, we anticipate that they will depend on the candidates' abilities, which in the current period are random variables. Note that the expected ability of the selected workers is $E(m / I_j > K) = \rho\,\lambda(K)$, where $\lambda(K)$ is the inverse Mills ratio of $\Phi(L/\rho)$, a positive and increasing function (in K). Given that candidates apply randomly, the average abilities of the workers hired by firm i, denoted m_i, are represented by a random variable with expected value of $\rho\lambda(K)$ in this Period 3. When we maintain a fixed value for the extensity of selection K, the intensity of the selection ρ will vary together with the expected ability of the selected workers, but it does not change the probability that a candidate is rejected.

Period 4: training provision. Between the period that the candidate is selected and training is provided, the firm can collect more information about the worker. For simplicity, we assume that in this period, the firm acquires perfect information about the workers' abilities. Firm i provides level of training t to a hired candidate j, thus obtaining the corresponding profits associated with the job position.

Our focus of interest is the relationship between the selection intensity of the candidates and their posterior training. Therefore, we assume that the profit of firm i when a candidate j is hired, denoted by π_i, is a function of the worker's abilities m and training (e.g., the time devoted to training), $t \geqslant 0$. This function can be written as

$$\pi_i = z_i q = z_i\left(\alpha t - \frac{t^2}{2} + \beta m t\right)$$

Let us interpret q as the increment in the workers' production associated with training and workers' ability. Hence, positive parameter z_i represents the firm-specific profit by-product and captures the fact that workers with the same training and ability can generate different levels of profits in different firms. Positive parameter α is the marginal productivity of the first (marginal) unit of training of a worker with the average ability $(m=0)$. We assume that this marginal productivity varies with the ability of the worker. In terms of Cunha and Heckman (2007) there are complementarities between abilities and training which are captured by a nonnegative parameter β. Furthermore, we assume that the marginal productivity is decreasing (in one unit of product[3]) with each (marginal) unit of training, what can be justified by the fact that firms train first the most productive aspects of the job. Note that the marginal productivity of training could be negative, which is interpreted as the costs of one extra unit of training being higher than its benefits. These costs and benefits (measured in terms of production) are not modeled here but are assumed to be included in parameters α and β.

Function π_i can be considered the second-order Taylor expansion of any functional form[4] that considers only workers' training and abilities. This is the minimum order of the Taylor expansion that represents the existence of complementarities among training and abilities in a single parameter and provides a simple expression for the optimal level of training, $t^* = \alpha + \beta m$, note that $t^* > 0$ for $m > -\alpha/\beta$ and $t^* = 0$ otherwise (see Appendix 1 for further details). When m is known, it is always possible to provide a level of training that generates nonnegative profits.

This benchmark is designed to compare the implications of the three contracts mentioned in the introduction. In restrictive contracts, the training decision is made in Period 1. Obviously, making decisions in advance has a cost because firms cannot adapt to future circumstances. To capture this cost, we assume that the marginal productivity of one unit of training is $\alpha = \alpha_2 + \epsilon_i$, where α_2 is a positive parameter denoting the expected marginal productivity of one unit of training at Period 1. In Period 4, between hiring and providing training to the worker, firms receive a specific shock ϵ_i with an expected value of zero, which encapsulates all the relevant information that might change the marginal productivity of training. The abilities of the candidate are unknown in Period 1 but revealed in Period 4, after the selection process (once the candidate is already working for the firm). The rest of the parameters of the profits' function are known in Period 1.

Types of contracts

The contracts that we will discuss are defined according to the moment when the training decisions are made and the information is used for making such decisions. In direct contracts (c=DI), training decisions are made in Period 4 and use all the information available about the parameters and abilities defined in the general framework (Period 1–4). In indirect contracts (c=IN), training decisions are also made in Period 4, but firms do not use the information available about the abilities of a particular worker to set the level of training. In restrictive contracts (c=R), training decisions are made in Period 1, before the selection process, thus firms do not use information either on I_j or on ϵ_i, Figure 1 summarizes this discussion. It shows the moment in which the exogenous variables are revealed and the endogenous variables (intensity and extensity of selection, and training) are implemented. Note that the main difference between the contracts analyzed is the moment at which the training decision is made.

Given that in our context, direct contracts make better use of the information available, they outperform (in terms of social surplus) indirect contracts, which in turn

Exogenous Variables:	Period 1 Abilities	Period 2 Information collection	Period 3 Candidates' selection	Period 4 Training provision
Candidate Abilities	m_j - (unknown)		I_j (Information with quality ρ)	m_j (known)
Profits of training	z_i - Profit by-product α_2-Training marginal productivity (expected) β - complementarities abilities and training			ε_i (realized)
Information costs	$C(\rho)$- Costs selection intensity			
Implementation of Endogenous variables:		ρ -selection intensity	K -selection extensity	t -training
Types of contracts: When *t* is decided?	Restrictive contracts			Direct contracts Indirect contracts

Figure 1. Timeline of the general framework.

outperform restrictive contracts, which make use of no information ($S_i^{DI} \geq S_i^{IN} \geq S_i^R$, see Appendices 1, 2, and 3 for formal details). If we assume that enforcing direct contracts is not possible (Chang & Wang, 1995; Kahn & Hubberman, 1988; Prasad & Tran, 2013; Prendergast, 1993), for example, due to the difficulties in demonstrating the abilities[5] of the worker to third parties, direct contracts would imply a low training effort of the workers ($t=0$ in our framework[6]) and consequently $S_i^{IN} \geq S_i^R \geq S_i^{DI} = 0$. Therefore, the use of different contracts can be justified by the inability to make training contracts contingent on the abilities of particular workers (this would justify the use of indirect contracts) and/or the shocks in the marginal productivity of training (both limitations would justify the use of restrictive contracts).

The data available does not include information about the specific training contract used by each firm; hence, we cannot investigate the contexts in which such contracts could be infeasible. Therefore, our research approach will be to deduce the empirical implications of each type of contract and infer which one fits the data best. In the empirical application, we do not have information about the firms' selection extensity; thus, henceforth, our discussion will focus on selection intensity. Next, we summarize the main implications of each type of contract.

Empirical implications and the role of education

In Barron et al. (1997), an exogenous level of training required for a position determines the level of selection intensity $(\partial\rho(\alpha_2,t)/\partial t>0)$ (Appendix 3 reproduces their main propositions). Empirical support for this relationship is provided by these authors and by an earlier study (Barron et al., 1989). These authors also find evidence of a positive association between selection intensity and workers' education. One way to reconcile these results with our framework is by interpreting education (e) as a proxy for the marginal productivity of training α_2. In terms of our model, this would imply that education directly and positively affects both the optimal level of training $(\partial t(\alpha_2)/\partial e>0)$ and selection intensity ($\partial\rho(\alpha_2,t)/\partial e>0$ when t is fixed), Table 1 summarizes these predictions.

On the contrary, several theoretical approaches suggest that education might be an indicator of candidates' abilities.[7] If this were the case, the percentage of educated workers in a job position would be an endogenous rather than an exogenous variable (e.g., α_2), as in the former case. In other words, although the candidates obtain education before applying to the position (Period 1), the candidates who are hired and, consequently, their level of education can be known only after the selection process (Period 4).

Our general framework assumes homogeneity in terms of the cost function of providing information $C(\rho)$, complementarities β and the expected marginal productivity of training α_2. Under these assumptions, and unlike the previous interpretation, education could be considered a proxy for workers' abilities.

Our empirical application is in line with this interpretation. It is based on firms' blue-collar workers, where differences in such parameters seem to be much less important than in data for several job positions. Furthermore, we will use variables as technological intensity that can help control for these factors. Therefore, we expect little variability in the expected marginal productivity of training (α_2). The analysis of direct and indirect contracts will be conducted assuming that the percentage of educated workers in a job position is an indicator of the average workers' abilities in that job position, m_i in terms of our model. This is the usual interpretation in the empirical literature (Heckman, 2000; Mincer, 1992; Rosen, 1976).

In our framework, the average ability of the workers hired by firm i, denoted m_i, is a random variable with expected value $\rho_i\lambda(K(\rho_i))$. For the empirical implementation, this can be represented as

Table 1. Summary of direct effects in restrictive contracts.

Exogenous variables	Endogenous variables	
	Training $t\ (\alpha_2, CV)$	Selection intensity $\rho\ (\alpha_2, t, CV)$
Training (t)		$\partial\rho / \partial t > 0$
Education (e $\approx$ training productivity α_2)	$\partial t / \partial e > 0$	$\partial\rho / \partial e > 0$

CV are other control variables, in our framework (β, z_i). See further details in Appendix 3.

Table 2. Summary of direct effects in indirect contracts.

Exogenous variables	Endogenous variables	
	Education e (ρ, CV)	Training $t(\rho, CV)$
Education (e $\approx$ average workers' ability m_i)		$\partial t / \partial e = 0$
Selection intensity (ρ)	$\partial e / \partial\rho > 0$	$\partial t / \partial\rho > 0$

CV are other control variables, in our framework (α_2, β, z_i). See further details in Appendix 2.

$$m_i = \rho_i \lambda\left(K\left(\rho_i\right)\right) + v_i$$

where v_i is a random variable with an expected value of zero. Note that the variance of this variable decreases as the number of the firms' workers increases.

The direct contract presented here tries to represent the most common approach in the empirical literature on firm-provided training (Kramer & Tamm, 2018), in which there is no information about the selection process. In this type of contract, the selection intensity is irrelevant for training decisions $(\partial t / \partial\rho = 0)$ (see Appendix 1 for a more formal discussion); in other words, $m_i = v_i$. The basic idea is that firms can perfectly tailor the provision of training to the workers' abilities in such a way that training always generates profits; thus, there is no reason to reject a candidate $(\partial e / \partial\rho = 0)$. In this case, the average level of training provided by a firm can be written as

$$t^* = \alpha_2 + \beta m_i + \epsilon_i = \alpha_2 + \beta v_i + \epsilon_i$$

Since we have assumed that education is a proxy of the workers' abilities, direct contracts predict that firms hiring (on average) more educated workers will provide them with more training $(\partial t / \partial e > 0)$. Table 2 summarizes the predictions of direct contracts.

On the contrary, the indirect contract assumes that the same level of training is provided to all workers in the same job position, but this level of training is decided once their abilities are known. Under this assumption, it is sensible for firms to reject the candidates that are expected to generate negative profits after training. Note also that those incentives will be higher for firms with higher unitary profit z_i and that this will introduce heterogeneity in selection intensity. This heterogeneity has been documented in large samples of firms (see, for example, Barron et al., 1989, 1997) and in-depth case studies (see, for example, Rivera, 2015, on elitist companies' selection processes).

In this contract, selection intensity positively affects selection extensity, $K(\rho_i)$ (see Appendix 2 for further details). Therefore, the average ability of the workers hired by firm i, denoted m_i, is a random variable with expected value $\rho_i \lambda(K(\rho_i))$. With these assumptions, the average level of training provided by a firm can be written as

$$t^* = \alpha_2 + \beta m_i + \epsilon_i = \alpha_2 + \beta\rho_i \lambda\left(K\left(\rho_i\right)\right) + \beta v_i + \epsilon_i$$

This contract predicts that the firms' selection intensity ρ positively affects the level of education $(\partial e / \partial\rho > 0)$ of the workers hired as well as their training $(\partial t / \partial\rho > 0)$. Note that if education would explain abilities only partially, controlling for selection intensity would reduce the positive effect of education on training. In fact, for a low variance of v_i (notice that in comparison with the population our sample is biased toward big firms), we might expect that the effect of education would be insignificant $(\partial t / \partial e = 0)$. Table 3 summarizes the predictions of this type of contract.

Finally, one could suggest that education always is a contractible variable. When education is perfectly related to workers' abilities, the optimal contract will be the direct one. In the case where education is unrelated to the workers' abilities, the optimal contract will be the indirect one. These contracts can be understood as the extreme cases of a combination of (direct) contracts based on education and (indirect contracts) with some (minimum) level of training associated with the job position.

Methods

Econometric specifications

We use cross-sectional data for estimating the following recursive triangular system of equations determining the education and training levels of workers

$$e_i = \alpha_1 + \lambda\rho_i + \lambda_x x_i + \epsilon_{i,1} \quad (1)$$

$$t_i = \alpha_2 + \beta e_i + \delta\rho_i + \beta_x x_i + \epsilon_{i,2} \quad (2)$$

where observations i are industrial plants, and the variables of interest are selection intensity ρ_i, education e_i, and training t_i. The other independent variables (x_i) will

Table 3. Summary of direct effects in direct contracts.

Exogenous variables	Endogenous variables	
	Education $e\ (CV)$	Training $t\ (\rho, CV)$
Education ($e \approx$ average workers' ability m_i)		$\partial t / \partial e > 0$
Selection intensity (ρ)	$\partial e / \partial \rho = 0$	$\partial t / \partial \rho = 0$

CV are other control variables, in our framework (α_2, β, z_i). See further details in Appendix 1.

be described in the next section. The parameters to be estimated are $\lambda = \partial e / \partial \rho$, $\beta = \partial t / \partial e$ and $\delta = \partial t / \partial \rho$ and β_x.

This specification allows us determine the type of contract that better fits the data by testing the following predictions (see Tables 1 to 3, above):

Direct contracts: $\lambda = 0, \beta > 0 \text{ and } \delta = 0$.

Indirect contracts: $\lambda > 0, \beta = 0 \text{ and } \delta > 0$.

Restrictive contracts:[8] $\lambda = 0, \beta > 0 \text{ and } \delta = 0$.

We do not claim that we provide evidence on causality among the variables. The econometric specification is chosen on the basis of firms' usually observed sequence of decisions, as described in the theoretical framework. A similar specification is used by Kramer and Tamm (2018) but with the difference that they do not analyze the impact of the intensity of firms' selection policies.

The estimations can be conducted with two alternative assumptions about error terms $\epsilon_{i,1}$ and $\epsilon_{i,2}$. If one assumes that they are uncorrelated with each other and with the independent variables, one could fit both equations individually (Kmenta, 1997, pp. 719–720). This is what has been assumed in most of the previous empirical literature.

If, on the contrary, one assumes that they are correlated with each other and with the independent variables (e.g., $\epsilon_{i,2}$ correlated with education or selection intensity or $\epsilon_{i,1}$ correlated with selection intensity), those equations would have to be estimated with methods that explicitly consider these facts. Given the cross-sectional nature of our data as well as the difficulty of finding appropriate instruments, our benchmark estimations will be conducted relying on the former assumption. For robustness, in Appendix 5, we also estimate equations (1) and (2) using instrumental variable methods. The similarity in results provides insights on the accuracy of our estimates.

Sample

The data are taken from a survey[9] of Spanish industrial plants, which was designed to obtain information on human resources and work organization practices for blue-collar workers. To ensure that the survey instrument adequately conveyed the intended research questions, the original questionnaire was fine-tuned with a pretest sample of 15 plant directors. The anonymity and 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 increases the possibility of obtaining information on some concepts, even if objective information is not available. The caveat is that it makes it difficult to complement the information with other sources and then search for instruments outside the sample.

The target group was a collection of manufacturing plants in mainland[10] Spain with 50 or more workers and whose economic activity was included in one of the 13 manufacturer sectors of the NACE[11] classification for 1993. The unit of observation was 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 was completed between 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 min long.[12] As some questionnaires were incomplete,[13] we ended up with 362 observations. Table 4 compares the distribution of the plants by size and economic sector among the population of Spanish manufacturing plants[14] and the sample. As seen, there are no important differences in terms of these variables. Note also that given the institutional structure of the Spanish state, we do not expect important differences across regions in the legislation regulating firms' economic activities.

Measures

Our survey provides measures for selection intensity as well as staff education and firm-provided training, which are the variables of interest in our theoretical discussion. Table 5 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 4.

The dependent variables in equations (1) and (2) are education attainment and investment in training of blue-collar workers, respectively. Regarding education, for

Table 4. Distribution of the plants by size and economic sector.

Variable	Category	% Sample	% Population
Size	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
Industry	Food, drink, and tobacco	17.13	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 nonmetallic mineral products	19.61	16.21
	Metallurgy metal equipment (excluding 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

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.

each plant, we have information on the percentages of blue-collar workers with no education (P_1, 7% on average), with primary education (P_2, 55% on average), with secondary education (P_3, 28% on average) and with university education (P_4, 10% on average). We aggregate all this information in the variable *Workers' education* ($P_1 + 2P_2 + 3P_3 + 4P_4$). This variable is the average of the standardized years of education of the plant's blue-collar workers when the differences between educational levels are held constant.[15] For robustness, we replicate the analyses (not presented but available upon request from the authors) by measuring education as the percentage of blue-collar workers with more than a primary education ($P_4 + P_3$). The main conclusions are maintained. The other dependent variable, *Training*, is measured by ordered (5-point Likert-type scale) answers to the question related to investments in blue-collar workers' training.

The independent variable selection intensity, *Selection* hereafter, is measured by ordered answers to the question about the variety of tools used during the selection process of blue-collar workers. To capture all available information, four dummies must be included in the estimations. A potential problem with this approach is that some of the categories have very few observations, which could cause collinearity in the estimations. 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 & Ben-Ner, 2017).

The survey allows us to control for a series of variables. *Job stability, Unions, High technological level* and *Limited workers' substitutability* are also measured by ordered (five-point Likert-type 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 5. Other control variables, *Collective agreements* and *Multinational*, are based on a binary answer from the questionnaire, so they do not need to be transformed. Finally, the control variable *Size* is measured by the logarithm of the number of workers, so it is continuous. For the parameters associated with these variables, we provide no prediction.

Results

Table 6 shows the results of the ordinary least squares (OLS) estimation of equation (1). The F-test rejects at the 1% level the null hypothesis that the parameter estimates of all explanatory variables are zero. The estimate on *Selection* is positive and statistically significant at the 1% level. Regarding the control variables, only *High technological level* is positively associated with *Workers' education*. The standard errors of this estimation are calculated using the White-Hubert procedure.[16] This procedure is

Table 5. Frequency distribution and descriptive statistics.

Variable [1]	Original coding [2]		Used coding [3]	
Training	*Likert scale*		Original coding	
	Nil or Very Low			0.03
	Low			0.12
	Average			0.56
	High			0.25
	Very High			0.04
Workers' education	*Percentages*		Standardized years	2.40 (0.43)
	% No studies	0.07 (0.16)		
	% Primary	0.55 (0.27)		
	% High school	0.28 (0.24)		
	% University	0.10 (0.13)		
Selection	*Likert scale*		Dummy	0.79
	Nil or very low	0.05		
	Low	0.16		
	Average	0.62		
	High	0.16		
	Very High	0.01		
Job stability	*Likert scale*		Dummy	0.51
	Nil or very low	0.02		
	Low	0.04		
	Average	0.44		
	High	0.43		
	Very High	0.08		
Unions	*Likert scale*		Dummy	0.32
	Nil or very low	0.06		
	Low	0.18		
	Average	0.44		
	High	0.25		
	Very High	0.07		
Collective agreements	*Dummy*		Original coding	0.57
Multinational	*Dummy*		Original coding	0.29
High technological level	*Likert scale*		Dummy	0.28
	Low	0.40		
	Middle low	0.32		
	Middle high	0.23		
	High	0.05		
Limited workers' substitutability	*Likert scale*		Dummy	0.62
	Total disagreement	0.05		
	Disagreement	0.21		
	Neither	0.12		
	Agreement	0.43		
	Total agreement	0.19		
Size	*Continuous*	201.21 (523.71)	Logarithmic	4.77 (0.78)
New ideas and initiatives (instrument for selection)	*Likert scale*		Dummy	0.69
	Total disagreement	0.04		
	Disagreement	0.07		
	Neither	0.21		
	Agreement	0.60		
	Total agreement	0.08		

(Continued)

Table 5. (Continued)

Variable [1]	Original coding [2]		Used coding [3]	
Education in the region (instrument for workers' education)	*Percentages*		Standardized years	2.70 (1.43)
	% No studies	0.01 (0.10)		
	% Primary	0.53 (0.49)		
	% High school	0.22 (45)		
	% University	0.24 (42)		

The number of observations is 362. Column [1] displays the variables used in the estimations. Column [2] presents the percentages of each category or variable mean and the standard deviations in parentheses. Column [3] presents the recodification of the original variables according to the criterion explained in the text. The unit of observation is the plant.

Table 6. Determinants of workers' education.

Independent variables	OLS
Selection	26.90***
	(4.94)
Job stability	–1.07
	(4.30)
Unions	–0.93
	(4.31)
Collective agreements	–3.98
	(4.35)
Multinational	2.78
	(4.99)
High technological level	23.38***
	(4.99)
Limited workers' substitutability	0.91
	(4.51)
Size	2.43
	(2.77)
Constant	202.05***
	(13.88)
R^2	0.15
F	9.69***
Observations	362

OLS: ordinary least squares.
Ordinary least squares (OLS) estimates (with robust standard errors). Standard deviations are in parentheses.
*$p < .10$, **$p < .05$, ***$p < .01$.

adequate under both the existence and inexistence of heteroscedasticity (in the latter case, the robust standard errors will be the conventional [OLS] standard errors). The Breush-Pagan and White tests cannot reject the null hypothesis of the existence of homoscedasticity.[17] The same methodology is used to calculate the standard errors for the estimation of equation (2).

Equation (2) is estimated by ordered probit probability models, since training is an ordered variable. We provide three different specifications. In Model 1, *Workers' education* is not included as an independent variable, Model 2 does not include *Selection*, and Model 3 includes both *Workers' education* and *Selection*. In this case, the Likelihood ratio tests (for homogeneity of variance) cannot reject the null hypotheses of the existence of heteroscedasticity.[18]

Columns [1] to [3] in Table 7 show the results. In the three alternative specifications, the Wald x^2 test rejects at the 1% level the null hypothesis that the parameter estimates on all explanatory variables are zero. The estimates on *Selection* (Models 1 and 3) are positive and statistically significantly different from zero at the 1% level. This is the variable with a higher z-value (5.52 and 5.16 in Model 1 and Model 3, respectively). Note that these results hold independently of whether we control for *Workers' education* and that the change in the coefficient across Models 1 and 3 is virtually negligible. The other relevant comparison is regarding the change in the estimate on *Workers' education* across Models 2 and 3. In Model 2, the estimate on *Workers' education* is positive, 0.03, and statistically significant at the 2.1% level. The introduction of the variable *Selection* in Model 3 reduces this estimate to 0.01, which is now statistically significant only at the 48% level. The Wald test of the null hypothesis that the *Workers' education* coefficient is equal across models reveals that the null can be rejected at the 1% significance level. Regarding the control variables, the statistically significant coefficients in the three models estimated are (ordered by z-values) those for *Size, Job stability, High technological level*, and *Limited workers' substitutability*, which are positive in all cases.

Appendix 5 provides additional estimations using instruments for education and selection intensity. The main results are similar to those shown above.

Discussion and managerial implications

Contributions to theory

This article extends the analysis of Barron et al. (1997) on the relationship between firms' training strategies and selection policies by relaxing the assumption that the training associated with a job position does not adapt to the

Table 7. Determinants of training.

Independent variables	Ordered probit		
	Model 1 Column [1]	Model 2 Column [2]	Model 3 Column [3]
Selection	0.98***		0.95***
	(0.18)		(0.18)
Workers' education		0.003**	0.001
		(0.001)	(0.001)
Job stability	0.40***	0.38***	0.40***
	(0.12)	(0.12)	(0.12)
Unions	–0.08	–0.08	–0.08
	(0.13)	(0.13)	(0.13)
Collective agreements	0.10	0.13	0.10
	(0.12)	(0.12)	(0.12)
Multinational	–0.02	0.01	–0.02
	(0.13)	(0.12)	(0.12)
High technological level	0.27**	0.24*	0.25*
	(0.13)	(0.13)	(0.14)
Limited workers' substitutability	0.22*	0.22*	0.22*
	(0.12)	(0.12)	(0.12)
Size	0.31***	0.33***	0.31***
	(0.08)	(0.08)	(0.08)
Cut1	0.48	0.77	0.69
Cut2	1.49	1.65	1.69
Cut3	3.34	3.38	3.55
Cut4	4.81	4.82	5.02
Pseudo R^2	0.11	0.06	0.11
Log likelihood	–369.14	–387.78	–368.91
R^2			
F			
Wald X^2	67.12***	49.64***	70.64***
X^2			
N	362	362	362

Columns [1] to [3] report ordered probit (robust standard errors) estimates. Standard deviations are in parentheses. The Wald test rejects at the 1% level the null hypothesis that the coefficients of *Workers' education* across Models 2 and 3 are the same (x^2 (1) is 13.37). *$p < .10$, **$p < .05$, ***$p < .01$.

information on workers' abilities gathered during the selection process and the period between the worker's hiring and when training is provided. When direct contracts (Heckman, 2000; Mincer, 1992; Rosen, 1976) are not feasible (Chang & Wang, 1995; Kahn & Hubberman, 1988; Prasad & Tran, 2013; Prendergast, 1993), firms might use indirect contracts, which provide the same level of training to all workers in the same position. Candidates are heterogeneous in terms of their abilities; therefore, the existence of complementarities between training and abilities implies that the expected profits of training will differ among candidates. For some of them, the cost of training does not compensate for its benefits. Our argument is that firms could take advantage of the information provided by the selection process (on the expected profits of training a particular candidate) to identify and reject those candidates. Therefore, the abilities of the selected workers will determine the profitability of their future training.

Contribution to empirics

Our theoretical framework is also an effort to encompass and compare the implications of the three training strategies considered in this study. The empirical application has identified clear patterns in the data suggesting the prevalent use of indirect contracts for blue-collar workers' positions in Spanish industrial plants. We found that plants collecting more information on candidates' abilities select those with more education and provide them with more training. In addition, when we control for selection intensity, the relationship between training and education is not statistically significant. Selection intensity is the variable with the greatest capacity to explain training. Surprisingly, the indirect contracts considered in this article have been virtually neglected by the existing empirical literature for interpreting their results.

Additional evidence is that both blue-collar workers' training and education levels are higher in technologically intensive sectors, whereas training is more prevalent in large firms and firms which offer more stability and find it harder to substitute workers.

Limitations and future research

As noted in the theoretical section, direct contracts will always outperform indirect contracts if they could both be implemented. However, in the context of training strategies, the incomplete contracts literature (Chang & Wang, 1995; Kahn & Hubberman, 1988; Prasad & Tran, 2013; Prendergast, 1993) suggests that, in general, it is not possible to make training contracts at the personal level. In the same line, the psychological literatures related to procedural justice (Folger & Cropanzano, 1998; Greenberg, 1996; Wiesenfeld et al., 2007) and social comparison (Ployhart et al., 2006; Taylor et al., 1990; Wood, 1989) also note the difficulties of implementing different human resources policies for workers in the same job position. One limitation of this study is that it does not provide evidence on these contractual difficulties. This evidence would allow us to better understand the reasons that explain the use of those contracts.

The cross-sectional data used in the empirical section allow us to present correlations conditional on a set of organizational characteristics but do not permit the testing of the causal order. Note, however, that the results of the instrumental variables methods presented as a robustness check (in Appendix 5) seem to suggest that endogeneity bias is not a severe problem. Future research with access to better-equipped data (in particular with data that allows for identification of exogenous shocks leading to different

selection intensity levels) and in other environments is needed to corroborate the evidence presented.

At a theoretical level, we do not explain the evidence related to the statistically significant control variables. Linking a significant control variable to one parameter of the model whose variations consistently explain this variable's relationships with the variables of interest is not straightforward. For example, one of these control variables is the firm's technological level. The evidence presented in the article, which is consistent with previous research (Acemoglu, 1997; González et al., 2016), shows a positive relationship between high technological level and blue-collar workers' training. We might explain this result, assuming that the marginal productivity of training (α_2) is higher in high-technology firms. However, this assumption fails to explain the findings that the technological level does not affect selection intensity or that it is positively related to workers' level of education—a relationship that has been documented in this and other studies (Bresnahan et al., 2002; Capozza & Divella, 2019; Moretti, 2004).

Tentative avenues for including these findings into the model could be to extend the model to incorporate new relationships. For example, in the model, we might include the increasing marginal productivity of workers' abilities (when there is no training). If this parameter is related to the firm's technological level, the model could explain the positive relationship between the technological level and workers' level of education reported in Table 6. This approach could also be useful for integrating into the theoretical framework other relationships found in the empirical application, such as the positive relationship among firm size, job stability, and substitutability of workers with training. We speculate that the degree of substitutability of workers could be a measure of the degree of specificity of workers' skills and that job stability could be a mechanism for providing incentives for firms to invest in specific skills. However, to avoid diverting attention from the model's central objective, we leave this exercise for future research.

The future research agenda should extend the model to include variables related to differences in the marginal productivity of abilities, economies of scale in the provision of training, and an analysis of the negotiation process for appropriating the profits of (general and specific) training among workers and firms. Furthermore, it could be interesting to analyze the implications of relaxing some of the model's other assumptions, such as, for example, that the information is always costly or that the quality of the information is independent of its realization.[19]

Managerial implications

At the current stage of knowledge, the theory suggests that managers should ask themselves whether they can implement training policies adapted to the particular workers allocated to the same job position. The evidence suggests that individualized training policies are not always implemented and indirect contracts are therefore used. In indirect contracts, all the workers in the same job position receive the same training, although the optimal level of training may not be profitable for some of them. One way to improve the firms' profits is to identify and reject those workers during the firms' selection processes. Therefore, firms should consider this when deciding the extensity and intensity of their selection policies. The article provides a formal model for guiding such decisions. Although establishing the training policies before selecting candidates (restrictive contracts) could provide some advantages, firms will lose the opportunity to generate profits related to the adaptation of training to the real abilities of the workers finally hired or changes in the relevant factors of their job. Updated workers' abilities comprise a key element of firms' competitiveness.

Conclusion

The article formalizes the link between a training strategy, which was (informally) suggested over the last three decades in the economics of private training literature, and the firms' selection policies. Our empirical analysis reveals that, at least in the context in which we conduct it, this is the prevailing strategy. This is surprising, given that the alternative strategies, in particular direct contracts, have been virtually the only ones used by empiricists to investigate the determinants of firm-provided training. To summarize, our analysis calls for an extension of the debate on the determinants of the training provision by firms and its links with selection policies.

Authors' Note

Eduardo C. Rodes is now affiliated with Pompeu Fabra University, Barcelona, Spain.

Acknowledgements

We thank participants at EALE 2018, Lyon, and EARIE 2018, Athens, for valuable comments.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Financial support from the Spanish Ministry of Economy, Industry and Competitiveness (R&D project: ECO 2017-86305-C4-2-R) is gratefully acknowledged. The usual disclaimer applies.

Notes

1. When an index X is negatively correlated, the simple transformation $I=-X$ is positively correlated. The assumption is just to avoid defining the information quality as the absolute value of ρ.
2. The social surplus function $S(.)$ is the sum of the welfare functions of the firm $B(.)$ and workers $U(.)$. We can write those functions as $B(.)=b(.)-w$ and $U(.)=u(.)+w$, with w being the wages paid by the firms to the workers. The social surplus function is $S(.)=b(.)+u(.)$. Throughout the article, we do not define $b(.)$, $u(.)$ (nor w), but it is quite simple to infer functions of this type compatible with our assumptions. Implicitly, we are assuming that the formulation of wages does not affect training decisions.
3. It is equivalent to assume that the decrease is o. Then, the profit by-product will be z_i/o, the marginal productivity of the first unit α/o and the complementarities parameter β/o.
4. $(\partial\pi(0,0))/\partial m$ is neglected for simplicity or assumed to be zero, $(\partial^2\pi(0,0))/\partial t^2=-z$; $(\partial\pi(0,0))/\partial t=z\alpha$; $(\partial^2\pi(0,0))/\partial t\partial m=\beta$.
5. It should be noted that in the related papers there are no differences between training (t) and abilities (m), which are aggregated into workers' skills (s): $s=\alpha t+\beta mt$. We are assuming that the contractual problems are related to the workers' skills before m, and then after the training process, $s=\alpha t+\beta mt$, but not with t.
6. A particular case of the profit function π_i is one where $z_i(\alpha t+\beta mt)$ are the profits for the firms and $z_i(t^2/2)$ are the costs for the workers of providing an effort or level of training effort t, which is a workers' decision based on the incentives provided by the contracts.
7. The analysis of why education is an indicator of workers' abilities is beyond the scope of this study. For example, see a discussion in Weiss (1995).
8. Strictly speaking, these authors propose the reverse causality (see Table 1). Unfortunately, with our data, it is impossible to provide evidence on the causal order.
9. 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.
10. 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.
11. The European Community statistical *classification* of economic activities.
12. Interviewer status was required by the questionnaire. Specifically, there are nine 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).
13. The missing values are spread throughout the different variables of the sample.
14. 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).
15. Under this assumption, the number of years of education follows the equation $Y=a+by$, 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, index y can be interpreted as a standardization of the years of education, $y=(Y-a)/b$, and the variable Education can be interpreted as the average of this index y for the plant workers.
16. We have also calculated the standard errors by clustering them at the level of randomization, that is, size and industrial sector (see Table 1), and the results (not presented) are very similar to the main results.
17. $X^2(1)=1.41$ (p-value=.23) and $X^2(46)=21.07$ (p-value=.98), for the Breush-Pagan and White tests, respectively.
18. $X^2(14)=34.70$ (p-value=.002), $X^2(14)=26.70$ (p-value=.02) and $X^2(14)=31.37$ (p-value=.005), for Models 1, 2, and 3, respectively.
19. Ortín-Ángel and Salas-Fumás (2007) assume that this realization (level of education) is related to the quality of the information about the workers' abilities.
20. http://database.espon.eu/db2/jsf/DicoSpatialUnits/DicoSpatialUnits_html/ch01s01.html
21. $x^2(8)=2.67$, so we cannot reject the null hypothesis of endogeneity.
22. The weak instrument test when there are more than one endogenous variable (Godfrey, 1999) suggests that the instrument for education is sufficiently correlated with the included endogenous regressor ($F=15.68$). However, it cannot reject the hypothesis that the instrument for selection is weak ($F=7.05$). Note, however, that this value is not very far from the conventionally accepted threshold.

References

Acemoglu, D. (1997). Training and innovation in an imperfect labour market. *The Review of Economic Studies*, *64*(3), 445–464.

Acemoglu, D., & Pischke, J.-S. (1998). Why do firms train? Theory and evidence. *Quarterly Journal of Economics*, *113*(1), 79–119.

Acemoglu, D., & Pischke, J.-S. (1999a). Beyond Becker: Training in imperfect labor markets. *Economic Journal*, *109*, F112–F142.

Acemoglu, D., & Pischke, J.-S. (1999b). The structure of wages and investment in general training. *Journal of Political Economy*, *107*(3), 539–572.

Barrenechea-Mendez, M. A., & Ben-Ner, A. (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, J. M., Berger, M. C., & Black, D. A. (1997). Employer search, training, and vacancy duration. *Economic Inquiry*, *35*(1), 167–192.

Barron, J. M., Black, D. A., & Loewenstein, M. A. (1989). Job matching and on-the-job training. *Journal of Labor Economics*, *7*(1), 1–19.

Bertrand, M., & Schoar, A. (2003). Managing with style: The effect of managers on firm policies. *The Quarterly Journal of Economics*, *118*(4), 1169–1208.

Bishop, J. (1991). On-the-job training of new hires. In D. Stem & J. Ritzen (Eds.), *Market failure in training* (pp. 61–96). Springer.

Boom, A. (2005). Firms' investments in general training and the skilled labour market. *Labour Economics*, *12*(6), 781–805.

Bresnahan, T. F., Brynjolfsson, E., & Hitt, L. M. (2002). Information technology, workplace organization, and the demand for skilled labor: Firm-level evidence. *The Quarterly Journal of Economics*, *117*(1), 339–376.

Capozza, C., & Divella, M. (2019). Human capital and firms' innovation: Evidence from emerging economies. *Economics of Innovation and New Technology*, *28*(7), 741–757.

Chang, C., & Wang, Y. (1995). A framework for understanding differences in labor turnover and human capital investment. *Journal of Economic Behavior and Organization*, *28*(1), 91–105.

Cunha, F., & Heckman, J. (2007). The technology of skill formation. *American Economic Review*, *97*(2), 31–47.

De Groot, M. H. (2005). *Optimal statistical decisions* (Vol. 82). John Wiley.

Folger, R., & Cropanzano, R. (1998). *Organizational justice and human resource management*. SAGE.

Frenette, M. (2004). Access to college and university: Does distance to school matter? *Canadian Public Policy/Analyse de Politiques*, *30*(4), 427–443.

Godfrey, L. (1999). Instrument relevance in multivariate linear models. *Review of Economics and Statistics*, *81*, 550–552.

González, X., Miles-Touya, D., & Pazó, C. (2016). R&D, worker training and innovation: Firm-level evidence. *Industry and Innovation*, *23*(8), 694–712.

Gordon, R. 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.

Greenberg, J. (1996). *The quest of justice on the job: Essays and experiments*. SAGE.

Heckman, J. J. (2000). Policies to foster human capital. *Research in Economics*, *54*(1), 3–56.

Kahn, C., & Hubberman, G. (1988). Two-sided uncertainty and up-or-out contracts. *Journal of Labor Economics*, *6*(4), 423–444.

Kmenta, J. (1997). *Elements of econometrics*. University of Michigan Press.

Kramer, A., & Tamm, M. (2018). Does learning trigger learning through adulthood? Evidence from training participation of the employed population. *Economics of Education Review*, *62*(1), 82–90.

Leuven, E. (2005). The economics of private sector training: A survey of the literature. *Journal of Economic Surveys*, *19*(1), 91–111.

Mincer, J. (1992). Human capital: A review. In C. Kerr & P. D. Standohar (Eds.), *Labor economics and industrial relations: Markets and institutions* (pp. 109–141). Department of Economics, Columbia University.

Moretti, E. (2004). Workers' education, spillovers, and productivity: Evidence from plant-level production functions. *American Economic Review*, *94*(3), 656–690.

Ortín-Ángel, P., & Salas-Fumás, V. (2007). Compensation dispersion between and within hierarchical levels. *Journal of Economics & Management Strategy*, *16*(1), 53–79.

Ployhart, R. E., Weekley, J. A., & Baughman, K. (2006). The structure and function of human capital emergence: A multilevel examination of the attraction-selection-attrition model. *Academy of Management Journal*, *49*(4), 661–678.

Prasad, S., & Tran, H. (2013). Work practices, incentives for skills, and training. *Labour Economics*, *23*, 66–76.

Prendergast, C. (1993). The role of promotion in inducing specific human acquisition. *Quarterly Journal of Economics*, *108*(2), 522–534.

Rivera, L. (2015). *Pedigree: How elite students get elite jobs*. Princeton University Press.

Rosen, S. (1976). A theory of life earnings. *Journal of Political Economy*, *84*(4, Part 2), S45–S67.

Schmidt, P. (1976). *Econometrics*. Marcel Dekker.

Taylor, S. E., Bram, B. P., & Aspinwall, L. G. (1990). Social comparison, stress and coping. *Personality and Social Psychology Bulletin*, *16*, 74–89.

Van den Steen, E. J. (2005). Organizational beliefs and managerial vision. *Journal of Law, Economics & Organization*, *21*(1), 256–283.

Van den Steen, E. J. (2018). Strategy and the strategist: How it matters who develops the strategy. *Management Science*, *64*(10), 4533–4551.

Weiss, A. (1995). Human capital vs. signalling explanations of wages. *Journal of Economic Perspectives*, *9*(4), 133–154.

Wiesenfeld, B. M., Swan, W. B., Jr., Brockner, J., & Bartel, C. A. (2007). Is more fairness always preferred? Self-esteem moderates reactions to procedural justice. *Academy of Management Journal*, *50*, 51235–51253.

Wood, J. V. (1989). Theory and research concerning social comparisons of personal attributes. *Psychological Bulletin*, *106*, 231–248.

Zellner, A., & Theil, H. (1962). Three-stage least squares: Simultaneous estimation of simultaneous equations. *Econometrica: Journal of the Econometric Society*, *30*(1), 54–78.

Appendix 1: direct contracts (DI)

Period 4: optimal training

$$t^* = argmax_t\, \pi_i = z_i\left(\alpha t - \frac{t^2}{2} + \beta m_j t\right) \quad \text{subject to } t \geq 0$$

For $m_j \geq -(\alpha/\beta)$, the first-order condition implies $t^* = \alpha - \beta m_j \geq 0$. The second-order condition $(\partial^2\pi/\partial t^2) = -1$ shows that an interior maximum exists.

Therefore, as $\alpha = \alpha_2 + \epsilon_i$, training increases (perfectly) with the marginal productivity of training (education): $(\partial t/\partial\alpha) = (\partial t/\partial\alpha_2) = (\partial t/\partial\epsilon_i) = 1$. Other static comparisons show that $(\partial t/\partial\beta) = m_j$ and $(\partial t/\partial m_j) = \beta > 0$.

For values of $m_j < -(\alpha/\beta)$ and $t > 0$, the profits are negative, $z_i\left(\alpha t - \frac{t^2}{2} + \beta m_j t\right) < z_i\left(-\frac{t^2}{2}\right) < 0$. Then, the optimal level of training is $t^* = 0$, and $p_i^{DI} = 0$.

Period 3: selection extensity

In any case, the profits will be nonnegative, so there is no reason to reject a candidate. Then, $\Phi(-\infty) = 0$, $K = L/\rho = -\infty$, $p_i^{DI} = E(\pi_i)(1 - \Phi(L/\rho)) = E(\pi_i)$.

$\pi_i = z_i((\alpha + \beta m_j)^2)/2$ when the optimal level of training is positive and $\pi_i = 0$ otherwise.

Period 2: selection intensity

$$Max_\rho : S(\rho) = p_i^{DI} - C(\rho) = E(\pi_i) - C(\rho)$$

Since $E(\pi_i)$ does not depend on the selection intensity, the firm is not going to invest, $C(\rho) = \rho = 0$.

Appendix 2: indirect contracts (IN)

Period 4: optimal training

The average ability of the workers hired by a firm m_i is a random variable with expected value $\rho\lambda(K)$. Let us write it as: $m_i = \rho\lambda(K) + v_i$, where v_i is a random variable with an expected value of zero.

$$t^*(\alpha_2, \epsilon_i, \beta, \rho, K) = argmax_t\, z_i\, \alpha t - \frac{t^2}{2} + \beta(\rho\,\lambda(K)t + v_i)$$
$$= \alpha + \beta(\rho\,\lambda(K) + v_i)$$

$\frac{\partial t}{\partial \alpha} = \frac{\partial t}{\partial \alpha_2} = \frac{\partial t}{\partial \epsilon_i} = 1$, training increases with its marginal productivity. Comparative statics show

$\frac{\partial t}{\partial \beta} = \rho\,\lambda(K) > 0$, higher complementarities increase training

$\frac{\partial t}{\partial \rho} = \beta\,\lambda(K) > 0$, $\frac{\partial t}{\partial K} = \beta\rho\lambda'(K) > 0$, increases in selection intensity and extensity increase training.

The inverse of the Mills ratio $R(K)$ is $\lambda(K) = \frac{\phi(K)}{(1-\Phi(K))} = \frac{1}{R(K)}$ and $\phi(.)$ is the probability distribution function (pdf) and of a normal distribution with a mean of 0 and a standard deviation of 1. For further details, see Gordon (1941).

Period 3: selection extensity

The expected value of training is: $E(t^*) = \alpha_2 + \beta(\rho\lambda(K))$

$$E(\pi_i) = \pi_i(\widehat{m_j}) = z_i(\alpha_2 + \beta\rho\,\lambda(K))$$
$$\left(\alpha_2 + \mathrm{E}(\epsilon_i) + \beta\widehat{m_j} - \frac{\alpha_2 + \mathrm{E}(\epsilon_i) + \beta\rho\lambda(K)}{2}\right)$$
$$\pi_i(L = \rho K) = z_i(\alpha_2 + \beta\rho\lambda(K))$$
$$\left(\alpha_2 + \mathrm{E}(\epsilon_i) + \beta\rho K - \frac{\alpha_2 + \mathrm{E}(\epsilon_i) + \beta\rho\lambda(K)}{2}\right) = 0$$

$M(K) = \lambda(K) - 2K = (\alpha_2/\beta\rho)$, where $M'(K) = \lambda'(K) - 2 < 0$ due that $0 < \lambda(K)(\lambda(K) - K) = \lambda'(K) < 1$

$$K = M^{-1}\left(\frac{\alpha_2}{\beta\rho}\right)$$
$$\frac{\partial K}{\partial \alpha_2} = \frac{1}{\beta\rho}M^{-1\prime}(K) < 0, \frac{\partial K}{\partial \beta} = -\frac{1}{\beta^2\rho}M^{-1\prime}(K) > 0,$$
$$\frac{\partial K}{\partial \rho} = -\frac{1}{\beta\rho^2}M^{-1\prime}(K) > 0$$
$$\frac{\partial \rho\lambda(K)}{\partial \rho} = \lambda(K) + \lambda'(K)\frac{\partial K}{\partial \rho} > 0$$

Period 2: selection intensity

$$Max_\rho : S(\rho) = p_i^{IN}(\rho) - C(\rho)$$
$$p_i^{IN} = z_i\left(\frac{\alpha^2 + (\beta\rho\lambda(K))^2 + 2\alpha\beta\rho\,\lambda(K)}{2}\right)(1 - \Phi(K))$$

Therefore, $\rho(z_i, \alpha_2, \beta)$.

Appendix 3: restrictive contracts (R)

Period 3: selection extensity

$$E(\pi_i) = \pi_i(\widehat{m_j}) = z_i t\left(\alpha_2 + \frac{t}{2} + \beta\widehat{m_j}\right)$$

Note that the minimum level of ability to obtain profits is L, $\pi_i(L) = z_i t\left(\alpha_2 + \frac{t}{2} + \beta L\right) = 0$. This implies $L = \frac{\frac{t}{2} - \alpha_2}{\beta} = \rho K$.

Therefore, $K(\alpha_2, \beta, \rho, t) = \frac{\frac{t}{2} - \alpha_2}{\rho\beta}$.

The static comparisons show: $\frac{\partial K}{\partial \alpha_2} = -\frac{1}{\rho\beta} < 0, \frac{\partial K}{\partial \beta} = \frac{\left(\alpha_2 - \frac{t}{2}\right)}{\rho\beta^2}, \frac{\partial K}{\partial \rho} = \frac{\left(\alpha_2 - \frac{t}{2}\right)}{\beta\rho^2}$ and $\frac{\partial K}{\partial t} = \frac{1}{2\rho\beta} > 0$.

Period 2: selection intensity

$$Max_\rho S = p_i^R(\rho, t) - C(\rho)$$

Since positive profits can be generated only when the worker is hired, the profit function can be written as

$$p_i^R(\rho,t)=z_i t\left(\alpha_2+\frac{t}{2}+\beta\rho\lambda\left(\frac{\frac{t}{2}-\alpha_2}{\rho\beta}\right)\right)\left(1-\Phi\left(\frac{\frac{t}{2}-\alpha_2}{\rho\beta}\right)\right)$$

Then, it can be shown that the optimal selection intensity, $\rho^*(\alpha_2,\beta,t,z_i)$, depends on a set of variables in the following way: $\frac{\partial\rho^*}{\partial\alpha_2}>0$, $\frac{\partial\rho^*}{\partial\beta}>0$, $\frac{\partial\rho^*}{\partial z_i}>0$, and $\frac{\partial\rho^*}{\partial t}>0$. See the problem solution below.

Period 1: training decision

The social surplus is maximized when

$$Max_t: p_i^R=z_i t\left(\alpha_2+\frac{t}{2}+\beta\rho^*\lambda\left(\frac{\frac{t}{2}-\alpha_2}{\rho^*\beta}\right)\right)\left(1-\Phi\left(\frac{\frac{t}{2}-\alpha_2}{\rho^*\beta}\right)\right)$$

$$t^*(\alpha_2,\beta,z_i)$$

The solution

The maximization problem

We can replace $t=2(\beta K\rho+\alpha_2)$ in $p_i^R(\rho,t)$

$$\begin{aligned}p_i^R(\rho,K)&=2z_i\beta\rho(K\beta\rho+\alpha_2)(\lambda(K)-K)(1-\Phi(K))\\&=2z_i\beta\rho(K\beta\rho+\alpha_2)(\phi(K)-K(1-\Phi(K))\end{aligned}$$

Note that when $t=0$, $p_i^R(\rho,K)=0$, $K>-\frac{\alpha_2}{\rho\beta}$.

Then, the problem of the firm can be rewritten as

$$\begin{aligned}&Max_{\rho,K}: S(\rho,K)\\&=2z_i\beta\rho(K\beta\rho+\alpha_2)(\phi(K)-K(1-\Phi(K))-C(\rho)\end{aligned}$$

The first-order conditions are

$$\begin{aligned}\frac{\partial S(\rho,K)}{\partial\rho}&=2z_i(2K\beta^2\rho+\beta\alpha_2)\\&(\phi(K)-K(1-\Phi(K))-C'(\rho)=0\end{aligned}\quad(3)$$

and

$$\frac{\partial S(\rho,K)}{\partial K}=\frac{\partial p_i^R(\rho,K)}{\partial K}=0\quad(4)$$

whereas the second-order conditions

$$\begin{aligned}&\partial^2 S(\rho,K)/\partial\rho^2=2z_i(2K\beta^2)\\&(\phi(K)-K(1-\Phi(K))-C''(\rho)<0\end{aligned}\quad(5)$$

and

$$\frac{\partial^2 p_i^R(\rho,K)}{\partial K^2}<0\quad(6)$$

The derivatives of the selection intensity. We can use the envelope theorem to find the effect on ρ^* of α_2, β, z_i, and t

$$\frac{\partial\rho^*}{\partial\alpha_2}=-\frac{\frac{\partial S(\rho,K)}{\partial\alpha_2 d\rho}}{\frac{\partial^2 S(\rho,K)}{\partial\rho^2}}=-\frac{\frac{\partial p_i^R(\rho,K)}{\partial\alpha_2 d\rho}}{\frac{\partial^2 S(\rho,K)}{\partial\rho^2}}$$

From the second-order condition (5), it follows that $\frac{\partial^2 S(\rho,K)}{\partial\rho^2}<0$. Therefore, we can establish that

$$\frac{\partial^2 p_i^R(.)}{\partial\alpha_2\partial\rho}=2z_i\beta(\phi(K)-K(1-\Phi(K))>0,\ \text{then}\ \frac{\partial\rho^*}{\partial\alpha_2}>0$$

$$\frac{\partial^2 p_i^R(.)}{\partial\rho\partial\beta}=2z_i(4K\beta\rho+\alpha_2)(\phi(K)-K(1-\Phi(K)),$$

$$\text{then}\ \frac{\partial\rho^*}{\partial\beta}>0\ \text{and}$$

$$\frac{\partial^2 p_i^R(.)}{\partial\rho\partial z_i}=2(2K\beta^2\rho+\beta\alpha_2)(\phi(K)-K(1-\Phi(K))>0,$$

$$\text{then}\ \frac{\partial\rho^*}{\partial z_i}>0$$

$$\begin{aligned}\frac{\partial p_i^R(.)}{\partial t\partial\rho}&=\frac{\partial p_i^R(\rho,K)}{\partial K\partial\rho}\frac{\partial K}{\partial t}\\&=z_i\phi(K)(\rho\beta(1-2KR(K))-\\&\alpha_2(R(K))/\rho+z_i\phi(K)(\beta(1-2KR(K))\end{aligned}$$

The first-order condition (see demonstrations below) is fulfilled when $\rho\beta(1-2KR(K))-\alpha_2(R(K))=0$. Given that $\alpha_2(R(K))>0$, then $(1-2KR(K))>0$ must occur. Therefore, in the optimum

$$\frac{\partial p_i^R(.)}{\partial t d\rho}=z_i\phi(K)(\beta(1-2KR(K))>0,\ \text{then}\ \frac{\partial\rho^*}{\partial t}>0$$

First-order condition (4)

$$0=\frac{dp_i^R(\rho,K)}{dK}=2z_i\beta\rho((\beta\begin{pmatrix}\phi(K)-K(1-\Phi(K))\\+(\rho\beta K+\alpha_2)(\phi'(K)\\-(1-\Phi(K))+K\phi(K)\end{pmatrix})$$

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

$$\frac{dp_i^R(\rho,K)}{dK}=2z_i\beta\rho\phi(K)\begin{pmatrix}\rho\beta(1-2KR(K))\\-\alpha_2(R(K))\end{pmatrix}=0$$

This is fulfilled when $0=\rho\beta(1-2KR(K))-\alpha_2(R(K))$, one can establish that $\frac{\alpha_2}{\rho\beta}=\frac{1}{R(K)}-2K=M(K)$.

It is easy to show that $M\left(-\frac{\alpha_2}{\rho\beta}\right)=\frac{1}{R\left(-\frac{\alpha_2}{\rho\beta}\right)}+2\frac{\alpha_2}{\rho\beta}>\frac{\alpha_2}{\rho\beta}$ and $\lim_{K\to\infty}M(K)=-K=-\infty$, as $\lim_{K\to\infty}1/R(K)=K$.

Given that $0<\lambda(K)(\lambda(K)-K)=\frac{1-KR(K)}{(R(K))^2}<1$, $\frac{dM(K)}{dK}=\frac{1-KR(K)}{(R(K))^2}-2<-1<0$, this guarantees the existence of a unique interior solution $K^*>-\frac{\alpha_2}{\rho\beta}$. Therefore, $K^*=K(\rho)=M^{-1}\left(\frac{\alpha_2}{\rho\beta}\right)$.

Second-order condition (6)

$$\frac{d^2p_i^R}{dK^2}=2z_i\beta\rho\phi'(K)(\rho\beta(1-2KR(K))-\alpha_2(R(K))+2z_i\beta\rho\phi(K)\begin{pmatrix}\rho\beta\begin{pmatrix}-2R(K)\\-2K(KR(K)-1)\end{pmatrix}\\-\alpha_2(KR(K)-1)\end{pmatrix}$$

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

$$\rho\beta(-2R(K)-2K(KR(K)-1))-\alpha_2(KR(K)-1)<0$$

Multiplying this expression by $1/\rho\beta R(K)$ and replacing $\frac{\alpha_2}{\rho\beta}=\frac{1}{R(K)}-2K$, the inequality above is equivalent to $\frac{1-KR(K)}{(R(K))^2}-2<-1<0$. QED.

Appendix 4: survey items used for the variables

Training	The investment in blue-collar workers' training, in hours and in money, is . . . Nil or very low, Low, Average, High, Very high
Workers' education	Please indicate the percentage of blue-collar workers in your plant who have the education listed. No studies, Primary education, High school, University
Selection	The variety of tools used during the selection process (interviews, personality and ability tests, simulations . . .) is . . . Nil or very low, Low, Average, High, Very high
Job stability	The commitment to indefinitely maintaining the employment relationship with our workers is . . . Nil or very low, Low, Average, High, Very high
Unions	How do you assess the influence of unions on your firm's blue-collar workers? Nil or very low, Low, Average, High, Very high
Collective agreements	Is there a plant- or firm-specific collective agreement that regulates the working conditions of your production workers? Yes, No.
Multinational	Does the parent company have any other production plants in foreign countries (outside of Spain)? Yes, No.
High technological level	Aeronautical industry Pharmaceutical industry Electronic, medical, optical, and computer equipment Chemical industry Machinery and metal equipment Electric equipment Motor vehicle and transport supply

(Continued)

Appendix 4. (Continued)

Low technological level	Rubber, plastic materials, and nonmetallic mineral products
	Metallurgy
	Food, drink and tobacco
	Textile industry, dressmaking, leather, and footwear
	Wood, cork, paper, and graphic arts
	Furniture and various manufacturing industries
Limited workers' substitutability	It is difficult to find in the market workers with the knowledge, attributes, and abilities of our workers, thus it is difficult to replace them with workers of the same value
	Total disagreement, Disagreement, Neither, Agreement, Total agreement
	Nil or very low, Low, Average, High, Very high
Size	Approximately how many workers did you have in 2005?
New ideas and initiatives (instrument for selection)	New ideas and initiatives are encouraged
	Total disagreement, Disagreement, Neither, Agreement, Total agreement

Appendix 5: endogeneity concerns

In this Appendix, we conduct the estimations using instrumental variable methods. To conduct the estimations, we need instrumental variables for selection intensity and education. Those variables need to be (1) correlated with the endogenous variable to be instrumented (relevance) and (2) uncorrelated with the error term of the equation for the dependent variable (independence). This is a challenge given the nature of the problem that we are studying.

Previous studies have used instruments for education (for a discussion of its relevance and independence, see Kramer & Tamm, 2018). One of these instruments is the percentage of the population with different levels of education (see Frenette, 2004). To construct this instrument, *Education in the region*, we collected information from the Spanish Institute of Statistics INE about the percentages of individuals with no education, with primary education, with secondary education, and with university education at the level of autonomous communities (NUTS2 Eurostat[20]) in 2007. Then, we aggregate this information in the same way as we did to create the variable *Workers' education*.

We are not aware of the use of instruments for selection intensity in the previous literature. The exogeneity in our theoretical model comes from the shock ϵ in the costs of collecting information about the workers' abilities for learning. We are unaware of a well-established theory on its determinants. However, it seems reasonable to assume the existence of economies of scope in the collection of information about different workers' capabilities, such as their learning abilities or capacity for initiative. In Spain, there is a public debate claiming that formal education does not stimulate the initiative capacity of students; therefore, we might expect the capacity for initiative to be unrelated to success in training or education. Following Van den Steen (2005), one could suggest that leaders who believe in the relevance of the initiative of their employees benefit more from matching with more proactive employees and thus will be more interested in collecting information about this particular capability. Bertrand and Schoar (2003) and Van den Steen (2018) are examples of empirical and theoretical papers suggesting that leadership style or beliefs are developed before and thus influence firms' policies.

Based on this discussion, we use as an instrument for *Selection* the dummy variable *New ideas and initiatives*, which is constructed on the basis of a survey containing 5-point Likert-type scale items asking interviewees whether the establishment encourages new ideas and initiatives. This question is in a different section of the questionnaire, that is, one related to the leadership style of plant managers, than the other questions, which are in the human resources section. There are methods for the estimation of this type of system when the independent variable is continuous (e.g., 3SLS), but we are not aware of such methods when it is an ordered variable. Therefore, we treat the ordinal variable *Training* as continuous. Before presenting the simultaneous estimations, we check that this change is innocuous. Therefore, we reproduce the estimations of Columns [1] to [3] using *Training* as a continuous variable. The results (not presented) are very similar.

To identify the model, we use the variables *New ideas and initiatives* and *Education in the region* as instruments for selection and workers' education, respectively. In Table 8, we show the results of simultaneous estimations of equations (1) and (2) using 3SLS (Zellner & Theil, 1962), which is consistent and the most efficient instrumental variable method among those that use only sample information about the systems of equations (Schmidt, 1976). The main conclusions identified in the benchmark models are maintained. The estimates on *Selection* are positive and statistically significant in both the *Workers' education* and *Training* equations, although they are now significant at the 3% and 4% levels, respectively. After *Selection* is controlled for, the estimate on *Workers' education* in the *Training* equation is nonsignificant at conventional levels.

We used alternative instrumental variable methods (e.g., 2SLS), and the conclusions were similar; thus, we do not present those estimations in the article, although they

are available upon request from the authors. At this point, we note that the results of the Hausman specification test[21] seem to suggest the existence of endogeneity. In this Appendix, we have tried to reduce the potential endogeneity using instrumental variables methods. However, as we have acknowledged above, the results of these estimations must be interpreted with caution. Our instruments for selection and education, although potentially theoretically justified, seem to be weak.[22] Therefore, in spite of our efforts, we cannot claim to prove causal effects.

Table 8. Determinants of training and workers' education.

Independent variables	3SLS		
	Selection (Eq. 0)	*Education* (Eq. 1)	*Training* (Eq. 2)
Selection		71.22**	2.61**
		(31.30)	(1.25)
Workers' education			–0.01
			(0.01)
Job stability	–0.01	–0.52	0.23
	(0.04)	(4.65)	(0.14)
Unions	–0.01	–2.24	–0.02
	(0.04)	(4.93)	(0.15)
Collective agreements	0.02	–2.53	0.005
	(0.04)	(4.72)	(0.14)
Multinational	0.04	–3.62	–0.07
	(0.05)	(5.51)	(0.16)
High technological level	0.05	17.42**	0.29
	(0.05)	(5.36)	(0.20)
Limited workers' substitutability	0.01	–2.37	0.11
	(0.04)	(4.87)	(0.15)
Size	0.04	1.30	0.12
	(0.03)	(3.56)	(0.10)
New ideas and initiatives (instrument for selection)	0.16***		
	(0.05)		
Education in the region (instrument for education)		0.07***	
		(0.02)	
Constant	0.46**	–4.14	2.55**
	(0.13)	(41.85)	(1.18)
X^2	22.14	68.93	20.28
N	362	362	362

Columns [1] to [3] report 2SLS estimates for equations (0), (1), and (2), respectively. Standard deviations are in parentheses.
*$p < .05$, **$p < .01$, ***$p < .001$.