The Role of Firm Size in Training Provision Decisions: the Spanish Case[•]

Laia Castany

Preliminary version

Please do not quote or cite without permission

(December 2006)

Universitat de Barcelona
AQR Research Group - IREA
Avda Diagonal 690, 08034 Barcelona
Tel: +34 93 4021011 FAX: +34 93 4021821
Email: Lcastany@ub.edu

Abstract.
Empirical evidence shows that small firms provide less training to their employees than their larger counterparts. Firms' provision of training is usually associated with certain characteristics such as having a more skilled labor force, using more complex technology, being more innovative, operating in more competitive markets, being participated by foreign capital or having less temporary workers, among others. Large firms are also more associated with these characteristics than small firms. Our hypothesis is that small firms provide less training because they are not associated with these characteristics with the same intensity as large firms. We estimate a two-part model, which considers firms' decisions on the provision of training as a double decision process, where, first, firms decide whether to provide training, and then, the quantity of it. We apply the Oaxaca-Blinder decomposition to analyze the differential in the provision of training by firm size. We obtain that small firms face more restrictions in their access to training and that the main reasons for them are related with their technological activity, the degree of competition that they face and the participation of foreign capital.

Keywords: Continuous training; Firm Size; Skilled Labor;
JEL codes: J21, L11, M53, L25

[•] Acknowledgements: This paper is part of my PhD dissertation and has been done during a research visiting at the University of Amsterdam under the supervision of Hessel Oosterbeek, with whom I am very grateful for his valuable suggestions and for discussing. I am also most grateful with Enrique Lopez-Bazo for his very helpful comments and support and to Elisabet Motellón for her advice on Yun's methodology.

1. Introduction

The National Reform Program for Spain in October 2005, framed in the Lisbon Strategy, highlights the necessity for Spain to increase and improve the quality of its human capital. The educational level of the Spanish labor force has experienced an important increase in the last decades. López-Bazo and Moreno (2007) show evidence of an important increase of human capital over the period 1964-2000. Concretely, the average years of education of the population in the private productive sector has increased from around 4 to 10 years. Nowadays, almost 100% of the 16 year-old population has received education. Although the educational level of the Spanish labor force has improved considerably in the last three decades in relation to other advanced economies, this economy is still far from them. For example, the percentage of population with university studies over the population aged between 15 and 64 was 88% of the average EU-15 in 2004 (Gual, Jódar and Ruiz, 2006).

However the qualification of the employees does not only depend on their schooling but also on their life-long learning, which includes continuous and occupational training. Training is distinguished from formal school and post school qualifications (which are viewed as formal education) and is generally defined as courses designed to help individuals develop skills that might be of use in their job.

The National Reform Program for Spain emphasizes the role of life-long learning as a key element for already occupied people to acquire knowledge and skills useful for their present and future employment, and for unoccupied people to reincorporate to the labor market. What is more, life-long learning also permits adapting workers' skills to the permanent evolution of job requirements and enhances the competitive position of workers and their employers. The main purpose of continuous training is to provide knowledge and adequate skills to occupied employees so that they could adapt to the changing requirements of firms at any moment. In this way, they become more competent and their professional performance is improved. Our study is focused on continuous training provided by the firms to their employees. De la Fuente and Ciccone (2003), among others, comment that there is clear evidence that training increases productivity at the firm level and that it is also a source of innovation and therefore long-term competitiveness of firms.

Spain has a very low percentage of population aged 25-64 receiving continuous training: in 2003, this percentage was around 25%, while the average EU25 is above

40% and Spain only performed better than Greece and Hungary. In 2004, 5.2% of the population received continuous training, while the average EU15 is 10.7% and the average EU25 is 9.9%.[1] Thus, a more intense effort regarding continuous training should be done, as it would help creating a more dynamic and competitive economy, as well as contribute to workers' social integration.[2]

Since December 1992, organizations of workers and firms, as well as the Spanish administration have signed different agreements to impulse continuous training (*Acuerdos Nacionales sobre Formación Continua*, ANFC). Since the II ANFC in 1997, training policies are particularly concerned with certain collectives of workers that face more difficulties in keeping its employment and/or more barriers to access training: that is the case of workers in SMEs, disabled workers, workers above 46-years-old, women and unqualified workers.

In this paper we analyze the underlying reasons that may cause small firms having more difficulties in accessing training. Training is generally found to be associated with certain firms' and employees' characteristics that determine firms' decision to invest in training and the quantity of it. Our hypothesis is that large firms are associated with such characteristics, which could partially explain the differences in training provision between small and large firms. Given that the Spanish economy is characterized for having a smaller firm size than other advanced economies, SMEs difficulties in access to training may constitute a limitation for the economy as a whole. As training is considered to increase workers' skills (and thus firms and workers are considered to become more productive), we are interested in analyzing what determines the firms' decision on the provision of training and how does it affect small and large firms' expenditure.

We use data drawn from the *Encuesta sobre Estrategias Empresariales* (ESEE), an unbalanced panel that collects information of a sample of Spanish manufacturing firms with at least 10 employees, and which is representative by firm size and industry. The question addressed in this paper is analyzed in the framework of a strand of literature that estimates the determinants of firm-related training. This approach analyzes the impact of different characteristics on firms' decision of whether providing some training or not. It is quite a common practice to use probit or tobit models to

[1] National Reform Program for Spain (2005, pp 36, 68), from the Lisbon Strategy in March 2000.

[2] The National Reform Program has the objective of increasing the percentage of population that received training from 5.2% in 2004 to 10% in 2008 and 12.5% in 2010.

estimate firms' decision to invest in training and the expenditure on it. A novelty of our study is considering the decision on the provision of training as a double-decision process. We argue that firms first decide whether they want to provide training or not and, after, they decide the amount they want to provide. We suggest estimating a bivariate sample selection model and a two-part model, which take this characteristic into account. The difference between the two models is that the former corrects for sample selection bias. We consider the strengths and weaknesses of the two models and discuss which could be more appropriate to model firms' decisions on training provision, both theoretically and empirically. Given the previous evidence of important heterogeneity among firms, we also control for firm-specific effects.

We apply the Oaxaca-Blinder decomposition to analyze the differential in the provision of training by firm size (the differential in the probability of providing training and the differential in the quantity). This methodology permits decomposing the individual effect of each variable in two components: the part of the differential due to differences in the levels of firms' characteristics and the part due to differences in the impact of these determinants on training.

We obtain that small firms face more restrictions in their access to training and that the main reasons for them are related with their technological activity, the degree of competition that they face and the participation of foreign capital. This can be seen as a limitation for small firms to become more competitive because they have a more restricted access to training, a tool that permits adapting the skills of their employees for becoming more competitive through taking more profit from their technological activity and their effort to compete in a foreign market.

In Section 2 we provide a revision of several theoretical arguments that explain differences in training provision by firm size. We also review different training determinants suggested in the literature. In Section 3, and departing from these determinants, we propose a specification and a model to estimate their impact. This model considers the provision of training as a double decision process, which can be considered as a novelty of our analysis. In Section 4 we offer evidence that small firms spend less on training than their larger counterparts. Next we show evidence that training is associated to certain firm characteristics and that, among firms with such characteristics, large firms provide more training. On the basis of this descriptive analysis, in Section 5 we provide the results of the estimation of our specification. First, we discuss whether it is appropriate to estimate a model that takes sample selection into account and, second, we introduce firm-specific effects. On the basis of the estimation for the small and large firms' subsamples, in Section 6 we apply the Oaxaca-Blinder decomposition to analyze the differential in the provision of training by firm size.

2. Determinants of Training

Continuous training of workers is an extension of the process by which human capital stock is enhanced by the school system; however, this part of the educational process takes place within firms. Some is formal and occurs in a structured environment, often in a classroom. Other is informal and involves supervision and work associated with the production process. The importance of analyzing continuous training at firm level, instead of employees' level, lies in the fact that decisions on the expenditure on training are made at firm level.

The empirical work by Black, Noel and Wang (1999) addresses the relationship between different training measures and firm size for a sample of US firms, and they find that large firms invest more in training. They argue that large firms have scale economies in the provision of both formal and informal training and more opportunities of doing co-worker training (i.e. if more than one person is doing the same task, then one of them can leave his or her job for a while to teach the new worker without interruption of the productive process). Baldwin, Gray and Johnson (1995) argue that large firms might have higher pay-off from their investment, and thus they would invest more. Holtmann and Idson (1991) argue that they face lower investment risks because they 'pool risks'. Barron, Black and Lowenstein (1987) argue that there are more possibilities of shirking in large firms, because when employers work cooperatively to produce a common output it is more difficult to disentangle the participation of each one. Then large firms will have higher monitoring costs. A way of reducing monitoring costs is training their employees. Also, according to Hashimoto (1979), large firms have access to cheaper capital to finance training.

On the other hand there is a strand of literature dedicated to explore the reasons why firms decide whether to train workers or the amount of training they provide to workers. Some relevant empirical works are Bartel (1989), Baldwin, Gray and Johnson (1995), Black and Lynch (1998), Blundell et al. (1999). For the Spanish case, see Alba-Ramírez (1994b), Peraita (2005) and Albert, García-Serrano and Hernanz (2005b). This literature estimates the impact of certain firm characteristics (determinants) that are supposed to be associated to training decisions.

In this paper we argue that large firms are often associated to some of these characteristics, while small firms are not, or not with the same intensity. If these characteristics are associated to higher training levels, they would explain in part why small and large firms follow different patterns in their training decisions, and thus the differences observed between them in terms of whether to provide training and on the amount of it. In the following paragraphs, we discuss how these characteristics may have an influence on the decisions on training and how they might differ by firm size.

First of all, training will be dedicated to those who have previously shown aptitudes to learn through a formal education process because they are supposed to be able to take higher profit from their expenditure on training (Black and Lynch, 1998; or Alba-Ramírez, 1994). So, firms with more qualified workers are likely to provide more training. Evans and Leighton (1989) find evidence of some sorting on ability characteristics across firm sizes. Zábojník and Bernhardt (2001) propose a model in which workers in larger firms and industries acquire more human capital. We argue that large firms are associated with a more qualified labor force, which could explain that they provide more training than their smaller counterparts.

The use of advanced and specialized technology requires specific knowledge and skills that are not easily found in the labor market and training is a way of acquiring such skills (Baldwin, Gray and Johnson, 1995). Technological changes occur at high speed and they require the continuous upgrading of the current labor force. There exists a wide debate on whether technological change leads to deskilling (technology permits separating tasks in other simpler tasks so that high skills are not so necessary) or leads to upskilling (technology makes the most repetitive tasks automatic so that workers are set free to perform tasks that require higher skills). However, the empirical evidence seems to favor more upskilling rather than deskilling. The skill-biased technological change effect has been mainly studied for the case of formal education, but a similar argument could be applied in the case of training (Osterman, 1995). Often, the innovative activity or the innovative effort of firms has also been included as a measure of its technological complexity. But we consider that the use of advanced technologies and the innovative activity and effort might both require separate training, as they are quite different processes.[3] When firms obtain an innovation, they will need to incorporate it in their production process. As before, the specific knowledge that the new process or product requires may not be easily found in the labor market. For example, when they launch a new product they may need train their sales workers. Or when they implement a process innovation, they may need to provide technical training to their production workers (Alba-Ramírez, 1994). Since Schumpeter (1942), different authors argue that large firms have an advantage over small companies as their financial situation allows them to be the most capable innovators. Huergo and Jaumandreu (2004b) find that innovation is narrowly related to firm size. We consider that the same arguments apply for a more intense use of advanced technologies in large firms. Thus, we expect that large firms innovate more and make a more intense use of advanced technologies and this can partially explain that they provide more training.

Investing in training is a way of increasing firms' competitiveness. So, firms exposed to more competitive markets may invest more in training as a strategy to make their employees more competitive and to be able to survive (see for example, Bartel, 1989). Small firms will be more vulnerable to highly competitive markets than large firms in the same market, so one would expect them to invest more in training. However the effect is not so clear because probably only large and very competitive firms will place themselves in such competitive markets were small firms could not survive (i.e. international markets).

Other authors argue also that foreign-owned firms are more likely to train workers. Very often, these firms are multinational firms, more efficient in their management, who employ more qualified workers and who have a more positive attitude toward workers' skills than domestic firms (see Görg and Strobl, 2004; Hughes, O´Connell and Williams, 2004).

Finally, firms with a high percentage of temporary workers are expected to invest less in training. This effect has an additional importance in the case of the Spanish labor market, as there is a high degree of temporary employment. See the works by Alba-Ramírez (1994) at firm level and Albert, García-Serrano, Hernanz (2005a) at employee level. On the one hand, if workers abandon the firm in the short term, the firm will not be interested in training them as it will not be able to capture the

[3] Baldwin, Gray and Johnson (1995) comment that the lack of available data on this question led many authors to use proxies such as the capital-labor ratio or some measure of productivity.

returns from such investment. On the other hand, the temporary worker does not have incentives to acquire the firm-specific human capital as he or she has a low probability of continuing in the firm. Oi (1983) finds that large firms have less rotation because of internal labor markets, thus we expect that they are more likely to provide training.

There are determinants of training for which we can not control. First, the percentage of unionized workers in the firm: it has been argued that unions bargain with the employer to achieve greater investment in training; also, quit rates tend to be lower in unionized firms, and thus, the costs of training employees are lower in unionized organizations (Wagar, 1997). Large firms tend to be more unionized and so they will be more likely to provide training. Second, due to the fact that we use a firm-level dataset, we can not have into account the personal characteristics of workers (workers' age, sex, experience, tenure, nationality, civil status, parents' education level; see Oosterbeek, 1996), as well as the workplace and personnel practices available (Total Quality Management, benchmarking, job sharing, self-managed teams, number of organizational levels, internal promotion, incentive-based retribution or joint decision-making; see Black and Lynch, 2000).

To summarize, large firms are usually associated to having a more qualified labor force. They are also associated to having a more intense technological activity (using more advanced technologies or having a higher innovative activity). Large firms are also associated with operating in more competitive markets (i.e. international) or being partially owned by foreign capital. Finally, large firms usually employ less temporary workers. There are theoretical arguments and some empirical evidence pointing to the fact that these characteristics of large firms could lead to a higher level of training in the firm.

3. Methodological Issues and Empirical Model

The ultimate purpose of this paper is analyzing whether small and large firms follow different patterns in their decisions of providing continuous training to their employees. We argue that large firms have certain characteristics that may determine a higher provision of training. It is a common practice to estimate a probit model to analyze what determines whether firms provide training to their employees or not. To analyze the determinants of firms' expenditure on training, it is also quite common to estimate a tobit model, which takes into account the fact that the dependent variable is censored at zero as, by nature, it can only take nonnegative values. See for example, Alba-Ramírez (1994), Black and Lynch (1998) or Black, Noel and Wang (1999). The MLE of the tobit model would provide consistent estimations if the error term is normal and homoskedastic.[4] Estimating the specification by OLS instead, would provide inconsistent estimates, as it assumes that the dependent variable can take both positive and negative values. Moreover, as the logarithm of zero does not exist, a common solution is to add a small positive constant but this constant is set arbitrarily. The main limitation of the tobit model is that it is quite a particular case, as it does not consider that the decision on the quantity of training may be a double-decision process: first, firms may decide whether to invest in training or not, and second, they decide the amount they will spend on it. It is especially the case when the decision on training and on quantity are motivated by different determinants, for example, when the decision on training involves incurring fixed costs, such as designing a training plan or evaluating the necessities on training of the firm. The fixed costs determine the decision on whether to spend or not, but they need not affect the decision on the quantity. Even in the case that the two decisions are depend on the same factors, the dependent variable may have observations that take value zero with a high frequency and this mass of zeros may respond differently to covariates than the observations with positive values. When this occurs, there are reasons to model the decision on training as two separate mechanisms, which is a generalization of the tobit model.[5]

Two-part models permit estimating the determinants of the expenditure on training (quantity equation) supposing that, in a first stage, the firm decides whether to provide training to its employees or not (participation equation). These models add flexibility in the sense that they allow that zeros and non-zeros are generated from different densities. There are two approaches to such flexible models: the sample selection model and the two-part model itself. The main difference between them is that the former takes into account a sample selectivity effect, which may cause biased estimations when it is omitted. In this paper, we estimate the two models and discuss which is preferred in the specific case of firms' training provision, both from a theoretical and applied perspective.

---

[4]Although heteroskedasticity can be modeled, the tobit is hypersensitive to extreme values in the distribution.

[5] See Cameron and Trivedi (2005, pp. 544-551) for a thorough explanation.

The most popular sample selection model is the bivariate sample selection model studied by Heckman (1979). The so-called heckit model comprises a participation equation, which may cause sample selection:

$$dTR_i^* = X_{1i}^{'}\beta_1 + \varepsilon_{1i} \tag{5.1}$$

where dTR_i^* is a censoring latent variable that reflects whether each *i*-firm would will to provide some training and X_{1i} is a vector of variables that determine this decision. The willingness of the firms to provide training cannot be observed, but we observe whether the firm spends some money on it. Define dTR_i as the censoring observed variable, which is a binary indicator that takes value 1 if we observe that the firm does some expenditure on training. So, $dTR_i = 1$ if $dTR_i^* > 0$ and $dTR_i = 0$ if $dTR_i^* \leq 0$.

Define TR_i as the firms' expenditure on training and $lnTR_i$ its logarithm, which is determined by a vector of variables X_{2i}. The quantity equation can be expressed as:

$$\ln TR_i = X_{2i}^{'}\beta_2 + \varepsilon_{2i} \tag{5.2}$$

where X_{2i} is a vector of variables that determine this decision. Assuming that the error terms ε_{1i} and ε_{2i} follow a bivariate normal distribution with zero means, standard deviation σ_1 and σ_2, covariance σ_{12} and correlation ρ:

$$E(\ln TR_i \mid dTR_i = 1) = X'_{2i}\beta_2 + \sigma_{12}\lambda_i\left(X'_{1i}\beta_1\right) \tag{5.3}$$

where $\lambda_i(X_{1i}^{'}\beta_1) = \phi(X_{1i}^{'}\beta_1)/\Phi(X_{1i}^{'}\beta_1)$ is defined as the inverse Mills' ratio, ϕ is the standard normal density function and Φ is the standard normal cumulative distribution function. The coefficients β_1 are obtained by first-step probit regression of dTR on X_1: $P(dTR = 1 \mid X) = \Phi(X_{1i}^{'}\beta_1)$. The heckit model augments the OLS regression on the quantity of training by the inverse Mills' ratio and then uses the positive values of TR to estimate the model by OLS. The estimate of β_2 is consistent, as it takes the sample selection bias into account.[6]

By introducing the inverse Mills' ratio, this model corrects for the possible sample selection effects. Sample selection appears when the error terms of the two equations are not independent, and thus the covariance of the error terms, σ_{12}, is different from zero. When σ_{12} equals zero, the heckit model simplifies to the two-part model, which simply uses the positive values of TR to estimate the model by OLS, obtaining consistent β_2 parameters. The two-part model was first proposed by Cragg

[6] The bivariate sample selection model can also be estimated by ML however it imposes stronger assumptions on the distribution of the error terms.

(1971). This model was especially designed for data on expenditure that contains a large number of zeros and a right-skewed distribution. The two-part model also departs from a participation and quantity equation. As before, the participation equation is estimated by a probit model and the quantity equation as least squares standard regression. The difference with the heckit model is that it does not include the inverse Mills' ratio term in the quantity equation to take into account possible sample selection:

$$E(\ln TR_i \mid dTR_i = 1, X'_{2i}) = X'_{2i}\,\beta_2 \qquad (5.4)$$

Departing from the discussion in Section 2, we include the following covariates in X_1 and X_2: the firm size (*SIZE*), the percentage of white collars (*WHITE*), the intensity of use of advanced technologies done by the firm (*ATLOW, ATMED, ATHIGH*), the innovative capacity of the firm (*INNOV*), the geographical scope of the firm market (*MARKET*), the foreign capital participation (*FOREIGNK*) and the percentage of temporary workers (*TEMP)*. As control variables, we include the intensity of use of the productive capacity (*PRODCAP*), a variable on whether the firm belongs to a group (*GROUP*) and finally a set of region (*DREG*), industry (*DSEC*) and year (*DYEAR*) dummies.[7]

All in all, we estimate the following equations as the quantity equations of these models. The former, corresponds to the heckit model and the latter to the two-part model:

$$\ln TR_i = X'_{2i}\,\beta_2 + \sigma_{12}\lambda_i\left(X'_{1i}\,\beta_1\right) + v_i \qquad (5.5)$$

$$\ln TR_i = X'_{2i}\,\beta_2 + \varepsilon_{2i} \quad , \quad TR_i > 0 \qquad (5.6)$$

In the following paragraphs we follow Dow and Norton (2003) in discussing which of the two models could be more appropriate in the case of the firms' provision of training. The choice between the two models is a controversial question and has led to intense debate over the years. First of all, we should consider carefully what kind of dependent variable we are trying to model. When analyzing continuous variables on expenditure on training with a large proportion of zeros, do we observe potential training-providers that for some reason did not decide to provide training to their employees? Or otherwise we observe firms that do not desire to provide any training to their workers? In other words, is there a latent positive expected training provision which might have been incurred under certain circumstances? These authors argue that

[7] See Appendix A5.1 for a more detailed explanation in the measurement of variables.

when the zeros do not represent zero values for the potential outcome, then the potential and observed outcome differ, and sample selection bias could appear.

Lynch (1993) argues that, in small firms, fixed costs of training are distributed across a smaller number of employees, and the production losses associated with a worker being away from the workplace can be higher in a small than in a large firm. Other fixed costs may be, for example, the design of a training plan or the evaluation of their necessities of training. We argue that some firms, in the presence of fixed costs, could obtain a low net benefit from their investment in training. Even thought they would will to provide some training to their employees, when the net benefit takes values below a certain level, we would observe a zero. If the firm obtained a net benefit above this level, they would decide to provide training and we would observe some positive value. According with this argument, a large mass of zeros includes potential training-providers that for some reason did not decide to provide training. Thus, we are interested in the potential outcome rather than the actual outcome.[8] By omitting the unobservable effect hidden in the potential outcome, we are only considering those firms that obtain a high net benefit from training, so that coefficients of the decision on the quantity of training would be biased. In this sense, we consider that such fixed costs could be hiding a latent expected training provision and thus causing a sample selection bias in the coefficients. In such case, the heckit model would be more appropriate while the two-part model would only be appropriate when sample selection does not exist.

Second, the heckit model may have problems of identification when the same regressors are included in the two equations, while in the case of the two-part model this is not a limitation.[9] The heckit model with normal errors is theoretically identified without any restriction on the regressors. However, if the same regressors are included in the two equations, this model is close to unidentified because $X_1=X_2$ leads to multicollinearity problems. As Cameron and Trivedi (2005, pp 551) comment, it can be very difficult to make defensible exclusion restrictions sometimes. In our case, it seems difficult to find at least one regressor that determines the decision on whether to provide training or not, but does not determine the quantity of training provided.

[8] Dow and Norton (2003) argue that labor economists "are generally interested in the potential wage. Observations without positive wage outcomes do not imply that an individual worked for zero wages". While in health economics, "researchers are interested in the public and private budgetary implications of actual expenditures" and "potential expenditures that are never incurred will not affect health care budgets".

[9] Although it is also possible to make exclusions in the case of the two-part model.

A test of $\sigma_{12}=0$ in the heckit can be used to test the null hypothesis that the two-part model is correct against the alternative hypothesis that the heckit is correct.[10] However, under collinearity between the covariates and the inverse Mills' ratio, the power of the t-test on the inverse Mills' ratio is limited and thus we can not use such test as a criterion to select between the two models; with low collinearity, the t-test is reliable. According with Leung and Yu (1996), the two models can be appropriate under different circumstances and the main sources of multicollinearity are imposing no exclusion restrictions, having low variability among regressors or a high degree of censoring. These authors recommend using the condition number to check for multicollinearity between the inverse Mills' ratio and the covariates in the quantity equation. The condition number is defined as the square root of the ratio of the largest to the smallest eigenvalue of the moment matrix $X'X$. Based on a series of Monte Carlo experiments, Belsley, Kuh and Welsch (1980) suggest that a condition number beyond 30 is indicative of collinearity problems.

Finally, using statistical criteria to select between the two models, Dow and Norton (2003) recommend the test proposed by Toro-Vizcarrondo and Wallace (1968), which they name an empirical MSE test. The original test statistic was derived for OLS models, but the intuition can be extended to the heckit and two-part model. This test consists on calculating the empirical MSE of both estimators, under the assumption that one model is consistent and correct. Then, the estimator with the lower empirical MSE is chosen. The empirical MSE for the supposed correct model will then involve only the variance component while that for the other model will involve its variance and its squared bias relative to the former.

The empirical studies that use firm-level datasets reveal the high degree of heterogeneity among firms with similar observed characteristics. This particularity of the data requires estimating a model that takes firm-specific effects into account. If there are significant unobserved time-invariant firm-specific effects that are correlated with the explanatory variables, then the simple pooled regression may produce biased and inconsistent estimates. The random effects model assumes that the individual heterogeneity is part of a compound error term and that the error term is uncorrelated

[10] Dow and Norton (2003) stress that if the coefficient of the inverse Mills' ratio is zero, the heckit reduces exactly to the two-part model, but the two-part model does not require the coefficient to be equal to zero. The models simply make different implicit distributional assumptions and they are only partially nested.

with the regressors.[11] In the case of micro-databases, where firms in the sample are selected randomly from a larger population, it is quite common to estimate a random effects model, rather than a fixed effects model (Cameron and Trivedi, 2005).[12]

4. The Dataset and Descriptive Analysis

We use a sample of Spanish manufacturing firms from the *Encuesta sobre Estrategias Empresariales* (ESEE), carried out by the *Fundación Empresa Pública* (FUNEP). This survey is an unbalanced panel that covers the period 1990-2002 and collects information on strategic decisions and behavior of firms. Every four years, firms answer a complete questionnaire (with those issues that are supposed to change yearly) and a reduced questionnaire, the rest of the years, so that full information is available in 1990, 1994, 1998 and 2002. The reference population of the ESEE is firms with 10 or more employees dedicated to one of the activities corresponding to divisions 15 to 37 from the CNAE-93, excluding division 23 (activities related to refinement of oil and fuel treatment). In the base time period, all the firms with more than 200 employees were required to participate (and so 70% of them did). The firms with 10 to 200 employees were sampled randomly by industry and four size strata, retaining about 5%, so that representativity for every industry and firm size was guaranteed. The ESEE is designed to change as industry composition evolves. Newly created firms are selected using the original selection criteria. There are also exits in the survey, due to death and attrition, and these firms have been replaced by others in their industry and size group so as to maintain representativity.

In this paper, we use information drawn from this survey that corresponds to years 2001 and 2002,[13] with 1515 and 1505 firms respectively. Out of these, 31.55% and 30.3% are large firms. For these firms, data is available for all the variables required. The ESEE considers that large firms are those with more than 200 employees, and small firms have between 10 and 200 employees.

---

[11] Although this is a quite strong assumption we prefer the random effects model rather than fixed effects because some of our variables do not change over time (for example, sector or region dummies) and our data does not has sufficient number of years to estimate a fixed effects model properly.

[12] Groot and Maassen van den Brink (2003) estimate a random effects probit model to analyze the frequency of training in Dutch firms. Barrios, Görg and Strobl (2003), Máñez et al. (2004), Licandro, Maroto and Puch (2004) among others also estimate a random effects model.

[13] The information on the firms' provision of continuous trainingin the ESEE is only available for 2001 and 2002.

Table 1 shows a descriptive analysis of training, both for the discrete variable (*dTR*) and for the expenditure per worker (*TR*) for the year 2001 and in relation with the other variables of interest. Table 2 shows the same analysis for 2002. First of all, one observes that 40.4% of the firms in the sample provided training in our period of analysis.[14] As we are interested in differences by size, we separate the total sample in the subsample of small and large firms: we obtain that 24% of small firms provide some training in 2001 and 25% in 2002; in the case of large firms it rises to 72% in 2001 and 78% in 2002. These values are quite in line with Alba-Ramírez (1994), who finds that, in 1988, around 60% of large firms provided training. The average real expenditure per worker and year is 39 euros in small firms in 2001 and 44.7 euros in 2002; in large firms, it rises to 130.7 euros in 2001 and 151.2 euros in 2002. We perform tests of equality of proportions and equality of means[15] that permit analysing whether the differences in the provision of training by size are statistically significant. We obtain that large firms provide more training and that the differences are significant at 1% for both 2001 and 2002.

As explained in Section 2, the provision of training is considered to be associated with certain firms characteristics or determinants. In Tables 1 and 2, we analyse whether the provision of training is associated with such characteristics. We split the total sample in two groups: firms with each characteristic and without it, and then we compare the proportion and the average expenditure on training per worker in the two groups.[16] Firms with a percentage of white collars above the median provide significantly more training. Innovative firms and firms that make a more intense use of advanced technologies provide more training (concretely, firms that make a high use of these tecnologies provide more training than firms with a medium use, and firms that make a medium use of advanced technologies provide more training than firms with a

[14] Data from the Eurostat (CVTS2) shows that, in 1999, the percentage of Spanish firms providing training by size class are the following: 10 to 49 employees: 23%; 50 to 249 employees: 49%; 250 employees or more: 80%. However, notice that these percentages refer to any sector, while we only take manufacturing firms into account. The sector that provides more training is the service sector, which is quite a large sector in the Spanish case.

[15] The t-test of equality of means used here in based on the approximation suggested by Welch, as explained in Ruiz-Maya and Martín-Pliego (1995). It assumes that variances for the two subsamples are different. This has previously been tested, and we obtain that the null of equal variances is rejected in all the cases.

[16] In the case where three categories are possible for a given variable, such as the case of the use of advanced technologies, we perform two comparisons: first we compare the training provision in firms that make a use of advanced technologies with a low intensity with those that use advanced technologies with a medium intensity; next, we compare the training provision in firms that make a use of advanced technologies with a medium intensity with those that use advanced technologies with a high intensity.

low use). Also, firms that operate in international markets and firms that are more participated by foreign capital, also provide significantly more training. Finally, those firms with a percentage of temporary workers below the median also provide more training. In all the cases, the tests of equality of proportions and equality of means reject the null that the two groups provide the same training at 1%. The only exception is the test of equality of proportions in the comparison between firms with a percentage of temporary workers above and below the median in 2001, which does not show any significant difference between the two groups. These findings show that training seems to be associated with these characteristics, as our a priori reasoning indicated.

In this paper, we argue that the difficulty in the access to training by small firms is associated with the above-mentioned characteristics. Our objective is analysing whether small and large firms follow different patterns in their training decisions in relation to these characteristics. So, we investigate whether small and large firms are also different after conditioning on these characteristics. Tables 1 and 2 show that, among those firms that have a percentage of white collars above the median, the large ones provide significantly more training than their smaller counterparts. Also, among those firms that have a percentage of white collars below the median, the large ones provide significantly more training. The same result is obtained for all the other characteristics and the tests of equality of proportions and means reject the null that small and large firms provide the same training at 1%. The only exception is the test of equality of means in the comparison between small and large firms with a participation of foreign capital above the median in 2001, which does not show any significant difference between the two groups. All in all, we observe a clear picture: firms with certain characteristics provide significantly more training, and among this group, large firms also provide significantly more training than small ones.

Table 1. Descriptive of training in 2001 by firms' characteristics and size

	dTR	Eq prop test (•)	Training/worker (euros)	Eq mean test (•)	# obs
Total sample	39.67		67.9427		1515
Small	24.59	17.6718***	39.0094	6.471***	1037
Large	72.38		130.7122		478
Low % white	24.14	12.3622***	24.5848	7.9201***	758
High % white	55.22		111.3578		757
Low % white - small	13.61	12.9419***	14.5179	5.1208***	595
Low % white - large	62.58		61.3323		163
High % white - small	39.37	10.3892***	71.9787	4.2366***	442
High % white - large	77.46		166.6136		315
Adv tech low	24.73	10.2389***	40.1577	3.5192***	845
Adv tech med	53.08		89.886		454
Adv tech high	69.91	4.1309***	130.5171	1.7606**	216
Adv tech low - small	17.16	11.7554***	25.6008	4.9264***	711
Adv tech low - large	64.93		117.3959		134
Adv tech medium - small	39.16	6.9742***	63.7311	2.2359***	263
Adv tech medium - large	72.25		125.9003		191
Adv tech high - small	47.62	4.5829***	87.1314	1.7752**	63
Adv tech high - large	79.09		148.3819		153
Non innovative	24.31	12.2155***	34.6759	6.0239***	757
Innovative	55.01		101.1656		758
Non innovative - small	16.26	10.7411***	22.1866	5.1567***	615
Non innovative - large	59.15		88.7666		142
Innovative - small	36.73	11.3397***	63.5259	3.9682***	422
Innovative - large	77.98		148.4392		336
National market	29.26	12.1025***	54.5181	4.058***	1032
International market	61.90		96.6263		483
National market - small	19.70	12.967***	33.7689	3.7217***	812
National market - large	64.55		131.1016		220
International market - small	42.22	8.3184***	57.9218	5.2698***	225
International market - large	79.07		130.3802		258
Low % foreign K	29.90	14.7036***	48.0743	7.0823***	1184
High % foreign K	74.62		139.0125		331
Low % foreign K - small	20.02	14.1695***	29.2313	3.9025***	929
Low % foreign K - large	65.88		116.722		255
High % foreign K - small	63.89	3.1229***	123.1185	0.9612	108
High % foreign K - large	79.82		146.7101		223
High % temp workers	38.71	0.7669	51.4381	2.9616***	757
Low % temp workers	40.63		84.4255		758
High % temp workers - small	23.19	13.2218***	31.1772	5.5807***	526
High % temp workers - large	74.03		97.5733		231
Low % temp workers - small	26.03	11.7764***	47.0714	4.5848***	511
Low % temp workers - large	70.85		161.7045		247

Note: (•) Supposing that the variances are different and unknown. It has been previously tested and the null that variances are equal cannot be rejected.
(***) (**) and (*) denote significant at 1%, 5% and 10%.

Table 2. Descriptive of training in 2002 by firms' characteristics and size

	dTR	Eq prop test (•)	Training/worker (euros)	Eq mean test (•)	# obs
Total sample	41.26		76.9683		1505
Small	25.26	19.1235***	44.6833	9.4218***	1049
Large	78.07		151.2381		456
Low % white	24.57	13.1639***	31.8816	9.1805***	753
High % white	57.98		122.1151		752
Low % white - small	13.66	13.3871***	18.4951	4.8865***	593
Low % white - large	65		81.4952		160
High % white - small	40.35	12.1556***	78.7395	6.6247***	456
High % white - large	85.14		188.9369		296
Adv tech low	24.94	10.8696***	47.4908	3.8328***	838
Adv tech med	55.31		91.0573		452
Adv tech high	75.35	4.9775***	162.2428	3.2984***	215
Adv tech low - small	16.97	12.7351***	30.8087	6.3557***	713
Adv tech low - large	70.4		142.6452		125
Adv tech medium - small	41.64	7.0894***	65.8023	3.1198***	269
Adv tech medium - large	75.41		128.1807		183
Adv tech high - small	47.76	6.3151***	107.5427	1.7135**	67
Adv tech high - large	87.84		187.0056		148
Non innovative	27.71	12.6221***	40.9176	7.6938***	877
Innovative	60.19		127.313		628
Non innovative - small	16.67	14.3266***	22.3721	7.1306***	696
Non innovative - large	70.17		112.2308		181
Innovative - small	42.21	10.4296***	88.6737	4.2921***	353
Innovative - large	83.27		176.9119		275
National market	30.25	12.4939***	52.7272	6.1329***	1015
International market	64.08		127.1822		490
National market - small	20.3	13.8469***	30.0197	6.7268***	813
National market - large	70.3		144.1191		202
International market - small	42.37	9.6549***	95.1983	2.7784***	236
International market - large	84.25		156.8996		254
Low % foreign K	31.67	14.5138***	52.7326	9.1928***	1184
High % foreign K	76.64		166.3612		321
Low % foreign K - small	21.46	15.0616***	35.4448	5.6428***	946
Low % foreign K - large	72.27		121.4481		238
High % foreign K - small	60.19	4.7851***	129.5345	2.2942***	103
High % foreign K - large	84.4		183.761		218
High % temp workers	37.33	3.0859***	59.8489	3.3923***	750
Low % temp workers	45.17		93.9744		755
High % temp workers - small	22.8	12.977***	40.9062	4.5742***	535
High % temp workers - large	73.49		106.9856		215
Low % temp workers - small	27.82	13.9855***	48.6148	8.5701***	514
Low % temp workers - large	82.16		190.7164		241

Note: (•) Supposing that the variances are different and unknown. It has been previously tested and the null that variances are equal cannot be rejected.
(***) (**) and (*) denote significant at 1%, 5% and 10%.

Tables 3 and 4 show the mean and standard deviation of the determinants of training for the total sample and small and large firms' subsamples. As we expected, large firms have more white collars, they are more innovative and use advanced technology with an intermediate and high intensity more than small firms do; large firms also operate more in international markets and they are more participated by foreign capital. As for small firms, they use advanced technology with low intensity more than large firms do and they have more temporary workers than large firms. Moreover, the differences in these characteristics between small and large firms are significant at 1% in all the cases.

These results suggest that large firms may provide more training because they are associated to such characteristics and this constitutes the point of departure for the remaining of our analysis. In the next section, we perform a causal analysis to see if such characteristics are driving the training decisions and if they have different influence in small and large firms. As we explain in Section 6.1, the differential in the provision of training could also be associated to a higher impact of these characteristics on the decisions of training.

Table 3. Descriptive of firms' characteristics by firm size in 2001

	Total sample		**Small**		**Large**		
	Mean	Std dev	Mean	Std dev	Mean	Std dev	Eq mean test (•)
Size	243.4686	699.7923	46.9967	46.3617	669.7058	1133.0016	12.0116***
% White collars	10.948	12.4904	9.4495	12.189	14.1988	12.5308	6.9146***
Adv tech low	0.5578	0.4968	0.6856	0.4645	0.2803	0.4496	16.1346***
Adv tech med	0.2997	0.4583	0.2536	0.4353	0.3996	0.4903	5.5743***
Adv tech high	0.1426	0.3498	0.0608	0.239	0.3201	0.467	11.4685***
Innovation	0.5003	0.5002	0.4069	0.4915	0.7029	0.4574	11.4287***
International market	0.3188	0.4662	0.217	0.4124	0.5397	0.4989	12.3346***
% Foreign K	19.3241	38.2651	8.6972	26.9353	42.3787	47.891	14.3646***
% Temporary workers	20.3932	22.7669	22.0143	24.9551	16.8763	16.5684	4.7403***
# obs	1515		1037		478		

Note: (•) Supposing that the variances are different and unknown. It has been previously tested and the null that variances are equal cannot be rejected.
(***) (**) and (*) denote significant at 1%, 5% and 10%.

Table 4. Descriptive of firms' characteristics by firm size in 2002

	Total sample		**Small**		**Large**		
	Mean	Std dev	Mean	Std dev	Mean	Std dev	Eq mean test (•)
Size	241.5015	697.9168	47.4211	47.155	687.9714	1148.1854	11.9087***
% White collars	11.6006	13.1976	9.9985	12.9629	15.2864	13.0066	7.2555***
Adv tech low	0.5568	0.4969	0.6797	0.4668	0.2741	0.4466	15.9687***
Adv tech med	0.3003	0.4586	0.2564	0.4369	0.4013	0.4907	5.4373***
Adv tech high	0.1429	0.35	0.0639	0.2446	0.3246	0.4687	11.2302***
Innovation	0.4173	0.4933	0.3365	0.4727	0.6031	0.4898	9.8046***
International market	0.3256	0.4687	0.225	0.4178	0.557	0.4973	12.4729***
% Foreign K	19.0452	38.1206	8.2364	26.3801	43.9101	48.0792	14.8993***
% Temporary workers	19.3388	22.1853	21.0709	24.416	15.3543	15.1985	5.5139***
# obs	1505		1049		456		

Note: (•) Supposing that the variances are different and unknown. It has been previously tested and the null that variances are equal cannot be rejected.
(***) (**) and (*) denote significant at 1%, 5% and 10%.

5. Estimation

The ultimate purpose of this study is to shed some light on the reasons why small firms provide less training than their larger counterparts. Using the Oaxaca-Blinder decomposition we intend to assess whether the difference in the probability of providing training and the levels of training between small and large firms are due to different levels in certain characteristics or determinants of training or whether it is due to the different impact of these characteristics on training in the two groups. The point of departure of the Oaxaca-Blinder decomposition is the estimation of auxiliary regressions for small and large firms separately. This methodology is applied on the basis of our preferred empirical specification. In the following subsections we select a specification out of different possibilities based on alternative definitions of innovative activity. We also discuss whether it is more appropriate a model that takes sample selection into account or not. Finally, we also introduce firm-specific effects and test whether the panel data estimations are more appropriate than the pooled data estimations.

5.1. The Two-Part Model vs. the Heckit Model

As commented in Section 4, around 60% of the observations of our dependent variable *TR* take value zero. This percentage indicates the existence of a high degree of censoring, and thus the necessity to consider that the zeros and positive observations may be generated from different processes. Departing from the arguments in Section 3, the main objective of this Section is selecting whether it is more appropriate to consider a model that takes the existence of sample selection into account or not.

Before focusing on this question, we consider whether the training provision is either contemporaneous to the innovation or it takes place some time after the innovation is obtained. The idea is that firms obtain process or product innovations and they intend to incorporate them to the production process as soon as possible. When workers need some training to adapt their skills to the requirements of the innovation, firms will have to provide training at the same time in which they obtain the innovation or some period after that. Given that firms are interested in recovering the returns of its innovative effort, they will try to incorporate the innovation as soon as possible. If firms provided training after obtaining the innovation, the new technology would be idle for a period of time. Thus, we expect that firms provide training to their employees in the same period they obtain the innovation. However, implementing a process innovation or launching a new product may take longer than simply adopting advanced technology. Thus, training could take place some time after the innovation is obtained.

In Table 5, we estimate specifications (5.5) and (5.6), defining the innovative activity as contemporaneous to the provision of training. In Table 6, we show the results when it is lagged one period. In columns (a), innovative activity is defined using two dummy variables (named *PRODUCT*, *PROCESS*) that take value one when the firm has obtained a product/process innovation. In columns (b), the innovative activity is defined using one dummy variable that takes value one when the firm has obtained a product or process innovation (*INNOV*). The first and second columns show the marginal effects[17] and coefficients of the participation equation. The participation equation is the same in the heckit and the two-part model.[18] The difference between the two models resides in the quantity equation, which, in the case of the heckit model, contains an additional term to account for sample selection. The third and fourth columns show the marginal effects and coefficients of the quantity equation in the heckit model. Finally, the fifth column shows the coefficients of the quantity equation in the two-part model.

[17] For each variable, the marginal effects of the probit model are calculated as the average across all the observations of the standard normal density multiplied by the coefficients obtained from the probit model estimation.

[18] To estimate the heckit, we use the STATA command `heckman.` To estimate the two-part model, we use the commands `dprobit` and `regress`.

In the participation equation, the innovative activity is positive and significant, for both definitions (a) and (b) in Tables 5 and 6. In the quantity equation, results are more diverse: when product and process innovations are contemporaneous, only the coefficient for process innovations is statistically different from zero; when they are lagged, only the coefficient for product innovations is significant; when the innovative activity is defined as a single dummy variable (*INNOV*) its coefficient is significant, both in the contemporaneous and lagged cases. These results seem to point out the different nature of product and process innovations in their effect on training provision. Actually, process innovations seem to have a contemporaneous effect on the quantity of training per employee, while product innovations seem to have an effect one period after the new product is obtained. This result may be explained by the type of training associated to each type of innovation. Even though analyzing the determinants of each kind of training would be a very interesting exercise, for the purposes of the present analysis we will simply consider product and process innovations defined as a single dummy variable that affects firms' training provision contemporaneously. In this view, we follow the approach by Alba-Ramírez (1994).[19]

[19] Given the particular behavior of product and process innovations, in Table A.1 in Appendix A.2 we repeat the same exercise but innovative activity is defined in the following way: product innovations are lagged and process innovations are contemporaneous. These variables are considered as two separate dummy variables (*PRODUCT, PROCESS*) and as one single dummy variable (*INNOV*). In this specification we obtain that the innovative activity is positive significant in the quantity equation.

Table 5. Estimation of two-part and heckit models. Contemporaneous product and process innovations

	(a) Product (t) // Process (t)					(b) Innov (t)				
	Participation eq		Quantity eq			Participation eq		Quantity eq		
			BSSM		two-part model			BSSM		two-part model
	mg eff	coef	mg eff	coef	coef	mg eff	coef	mg eff	coef	coef
size	0.1279***	0.3395***	0.5116	0.0606	-0.0193	0.1318***	0.3507***	0.5213	0.0812	-0.0246
	(0.0108)	(0.0288)		(0.0697)	(0.0396)	(0.0107)	(0.0285)		(0.0752)	(0.0395)
white	0.0053***	0.0142***	0.029	0.0233***	0.0203***	0.0053***	0.0142***	0.0285	0.0236***	0.0198***
	(0.0009)	(0.0024)		(0.0039)	(0.0033)	(0.0009)	(0.0024)		(0.004)	(0.0033)
atmed	0.1297***	0.3387***	0.5278	0.0849	-0.0079	0.1324***	0.3463***	0.5373	0.1167	-0.0043
	(0.0251)	(0.0652)		(0.1116)	(0.0907)	(0.025)	(0.065)		(0.1158)	(0.0898)
athigh	0.1448***	0.3726***	0.6377	0.1965*	0.114	0.1511***	0.3893***	0.6592	0.2274*	0.1164
	(0.0368)	(0.0932)		(0.1224)	(0.1101)	(0.0366)	(0.0927)		(0.1271)	(0.1094)
product	0.1119***	0.2912***	0.4348	0.0251	-0.0293					
	(0.0274)	(0.0704)		(0.0919)	(0.0799)	0.1668***	0.4411***	0.7456	0.3209***	0. 1935***
process	0.1465***	0.3827***	0.6649	0.2585***	0.1767**	(0.022)	(0.0585)		(0.1104)	(0.0806)
	(0.0244)	(0.0634)		(0.0998)	(0.0774)					
market	0.1094***	0.2866***	0.4809	0.1598	0.0865	0.1109***	0.2912***	0.48	0.1719*	0.0792
	(0.0245)	(0.0637)		(0.0957)	(0.082)	(0.0244)	(0.0634)		(0.098)	(0.0819)
foreignk	0.0011***	0.0028***	0.0051	0.0027***	0.0022***	0.001***	0.0027***	0.0049	0.0029***	0.0023***
	(0.0003)	(0.0009)		(0.001)	(0.0009)	(0.0003)	(0.0009)		(0.001)	(0.0009)
temp	-0.0002	-0.0005	-0.003	-0.0063***	-0.0059**	-0.0002	-0.0005	-0.0032	-0.0065***	-0.006**
	(0.0006)	(0.0015)		(0.0024)	(0.0031)	(0.0006)	(0.0015)		(0.0025)	(0.0031)
prodcap	-0.0012	-0.0031	-0.0048	-0.0009	-0.0003	-0.0012	-0.0032	-0.0049	-0.0009	0
	(0.0008)	(0.0021)		(0.003)	(0.0029)	(0.0008	(0.0021)		(0.003)	(0.0029)
group	0.0377	0.0996	0.1985	0.1445	0.1121	0.0361	0.0958	0.1955	0.1624	0.1218
	(0.0301)	(0.0791)		(0.102)	(0.099)	(0.0299	(0.079)		(0.1032)	(0.0991)
dyear	-0.0311	-0.0827	-0.2192	-0.2721***	-0.2554***	-0.0307	-0.0816	-0.2142	-0.2731***	-0.2525***
	(0.021)	(0.0557)		(0.0729)	(0.0719)	(0.0208	(0.0555)		(0.0735)	(0.0719)
cons		-1.9872***		3.1591***	3.9768***		-2.0382***		2.8695***	3.9481***
		(0.4104		(0.8472)	(0.6068)		(0.4101)		(0.8945)	(0.605)
H0:dsec=0	Yes		Yes		Yes	Yes		Yes		yes
	46.37***		55.05***		2.97***	47.80***		56.89***		3.11***
H0:dreg=0	Yes		Yes		Yes	Yes		Yes		yes
	65.12***		30.79***		1.87**	67.08***		30.48***		1.83**
num obs	3020		1222		1222	3020		1222		1222
Pseudo R	0.3469		-		0.1781	0.3431		-		0.1789
pseudolnL	-1331.1115		-		-	-1338.8303		-		-
rho	-		0.353		-	-		0.4425		-
sigma2	-		1.2649		-	-		1.2888		-
sigma12	-		0.4465		-	-		0.5703		-
			(0.3268)					(0.3484)		

Note: Standard deviation in parentheses
(***) (**) and (*) denote significant at 1%, 5% and 10%.

Table 6. Estimation of two-part and heckit models. Lagged product and process innovations

	(a) Product (t-1) // Process (t-1)					(b) Innov (t-1)				
	Participation eq		Quantity eq			Participation eq		Quantity eq		
			BSSM		two-part model			BSSM		two-part model
	mg eff	coef	mg eff	coef	coef	mg eff	coef	mg eff	coef	coef
size	0.1275***	0.3389***	0.5092	0.0575	-0.0212	0.1306***	0.3478***	0.5208	0.0665	-0.0196
	(0.0108)	(0.0288		(0.0714)	(0.0393)	(0.0107)	(0.0286)		(0.075)	(0.0388)
white	0.0053***	0.0141***	0.0287	0.0228***	0.0198***	0.0053***	0.0142***	0.0287	0.0231***	0.0199***
	(0.0009)	(0.0024		(0.0039)	(0.0033)	(0.0009)	(0.0024)		(0.004)	(0.0033)
atmed	0.1279***	0.3344***	0.5227	0.0881	-0.0047	0.1307***	0.3422***	0.5323	0.095	-0.0043
	(0.0251)	(0.0652		(0.1139)	(0.0909)	(0.025)	(0.065)		(0.116)	(0.0909)
athigh	0.1499***	0.386***	0.6576	0.1961	0.1104	0.1554***	0.4005***	0.6799	0.209*	0.115
	(0.0368)	(0.0931		(0.1254)	(0.1099)	(0.0367)	-0.0929		(0.1281)	(0.1103)
product	0.1028***	0.2687***	0.4718	0.1954**	0.1414*					
	(0.0262)	(0.0677		(0.0904)	(0.0814)	0.1702***	0.454***	0.7456	0.2631**	0.153**
process	0.1293***	0.3401***	0.531	0.0959	0.0255	(0.0216)	(0.0583)		(0.1148)	(0.0785)
	(0.0234)	(0.0613		(0.0952)	(0.0782)					
market	0.1171***	0.3072***	0.5108	0.1595*	0.0852	0.1174***	0.3083***	0.5127	0.1708*	0.0916
	(0.0244)	(0.0634		(0.0977)	(0.0826)	(0.0243)	(0.0632)		(0.0993)	(0.0826)
foreignk	0.0011***	0.0028***	0.0051	0.0028***	0.0022***	0.001***	0.0027***	0.0049	0.0028***	0.0022***
	(0.0003)	(0.0009		(0.001)	(0.0009)	(0.0003)	(0.0009)		(0.001)	(0.0009)
temp	-0.0002	-0.0005	-0.0031	-0.0063***	-0.0059**	-0.0002	-0.0005	-0.003	-0.0065***	-0.006
	(0.0006)	(0.0015		(0.0025)	(0.003)	(0.0006)	(0.0015)		(0.0025)	(0.0031)
prodcap	-0.001	-0.0026	-0.004	-0.0005	0.0001	-0.0011	-0.003	-0.0045	-0.0009	-0.0002
	(0.0008)	(0.0021		(0.003)	(0.0029)	(0.0008)	(0.0021)		(0.003)	(0.0029)
group	0.0369)	0.0977	0.196	0.146	0.1133	0.0357	0.0947	0.1908	0.148	0.1143
	(0.03)	(0.079		(0.1025)	(0.1001)	(0.0299)	(0.079)		(0.1026)	(0.1001)
dyear	-0.0428**	-0.1139**	-0.2666	-0.28***	-0.2584***	-0.0421**	-0.1122**	-0.2636	-0.2842***	-0.26***
	(0.021)	(0.0559		(0.074)	(0.072)	(0.0209)	(0.0558)		(0.0746)	(0.0719)
cons		-2.0569***		3.1089***	3.937***		-2.0382***		2.8695***	3.9481***
		(0.4117)		(0.8758)	(0.6119)		(0.4101)		(0.8945)	(0.605)
H0:dsec=0	yes		Yes		yes	yes		yes		Yes
	46.20***		57.64***		3.16***	46.60***		56.25***		3.14***
H0:dreg=0	yes		Yes		yes	yes		yes		Yes
	64.97***		29.55***		1.84***	66.74***		30.50***		1.88**
num obs	3020		1222		1222	3020		1222		1222
Pseudo R	0.3442		-		0.1772	0.344		-		0.1772
pseudolnL	-1336.4628		-		-	-1336.9093		-		-
rho	-		0.3493		-	-		0.37		-
sigma2	-		1.2651		-	-		1.2703		-
sigma12	-		0.4419		-	-		0.47		-
			(0.3403)					(0.3541)		

Note: Standard deviation in parentheses
(***) (**) and (*) denote significant at 1%, 5% and 10%.

Departing from the results in Table 5 columns (b), our preferred specification, we are interested in selecting whether it is more appropriate a heckit model or a two-part model. According with Section 3, we discuss whether the zeros observed in the dependent reflect that firms are not interested in providing training (actual outcome) or otherwise they hide some latent expected training provision that only becomes positive under certain circumstances (potential outcome). We argue that, in the presence of fixed costs (Lynch, 1993), some firms can not afford to provide training and we observe a zero in the variable measuring the expenditures on training. If the fixed costs were smaller, they would decide to provide training and we would observe some positive value. In this sense, such fixed costs can be hiding a latent expected training provision. From this perspective, we are interested in the potential outcome and the heckit model seems to be more appropriate.

Next, we are interested in analyzing whether, in practice, sample selection exists for the case of firms' provision of training. The t-test on the inverse Mills' ratio is used to test the null that the two-part model is correct against the alternative that the heckit is correct. When the same regressors are included in the two equations of the heckit model, multicollinearity problems arise and the model is close to unidentified. However, in our empirical specification, it seems difficult to find at least one regressor that can be included in the participation equation but not in the quantity equation. When collinearity problems appear, the t-test on the inverse Mills' ratio is not an appropriate tool to select between the two models. For the total sample, the condition number for the covariates is 26.9, and after including the inverse Mills' ratio it takes a value of 36.9. As suggested in Cameron and Trivedi (2005, pp 554), although the condition number including the inverse Mills' ratio takes a value above 30, the increase when including this regressor is very small, for which we do not consider that multicollinearity problems are severe. In such case, the t-test on the inverse Mills' ratio can be considered a useful tool to select between the two models. Table 5 columns (b) show that the coefficient of the inverse Mills' ratio takes value 0.5703 and it is not statistically significant. Thus, the null that the two-part model is correct cannot be rejected for the total sample. In the case of the subsample of small and large firms, we obtain similar results.[20] Although from a theoretical point of view, we argue that sample

[20] For the subsample of small firms the condition numbers are 23.3 for the covariates and 35.8 after including the inverse Mills' ratio. For the subsample of large firms, the condition numbers are 41.2 and 72.2, respectively. In the case of small firms, multicollinearity problems can not be considered severe, while some more difficulties appear in the case of large firms. Table A.2 at the Appendix A.2 offers the results of the estimation of the two-part model and bivariate sample selection model for the total sample and the small and large firms' subsamples corresponding to the preferred specification (*INNOV* contemporaneous).

selection could exist, a reliable significance test on the inverse Mills' ratio obtains that in practice it is more appropriate to consider a two-part model.

Finally, to obtain further evidence on which of the two models seems more appropriate, we use statistical criteria to choose between them. As explained in Section 3, first we consider that the two-part model is the "true" model, and next, the heckit model. We select the model with smaller empirical mean squared error under the two assumptions. Table A.3 at the Appendix A.2 offers the results for these tests. For most of the variables of interest in our empirical specification, we obtain that the MSE for the two-part model is smaller than the MSE for the heckit model, indicating that the former is more appropriate. The only exception is the variable on the percentage of temporary workers, for which the model that accounts for sample selection seems more appropriate. As for the control variables, the same result is obtained and the two-part model is preferred with the exception of some regional dummies. Under the two assumptions, the results are similar, indicating the robustness of the results of the empirical MSE analysis. Thus, as obtained through the test on the inverse Mills' ratio, the two-part model seems to be more appropriate to model the firms' decision on the provision of training. The same result is obtained when applying the empirical MSE test for the small and large firms' subsamples.

The results for the estimation of the two-part model are shown also on Table 5 columns (b). The first and second columns show the marginal effects and coefficients of the probit corresponding to the participation equation. The fifth column shows the coefficients of the OLS estimation of the quantity equation. In the participation equation for the total sample, almost all the variables of interest are significant, except the percentage of temporary workers, and have the expected sign. In the quantity equation, only *WHITE, INNOV*, *FOREIGNK* and *TEMP* are significant. The results for the subsample of small and large firms are shown on Table A.2 at the Appendix A5.2. Results show the existence of certain differences in the behavior of small and large firms in their decisions on the quantity of training. In Section 6 we use the Oaxaca-Blinder decomposition to investigate further the underlying reasons of such differences.

5.2. The Two-Part Model with Random Effects

The empirical evidence highlights the existence of high heterogeneity among firms with similar characteristics. The random effects model permits taking unobservable characteristics of the firms into account. In this Section we estimate the participation and quantity equations introducing a firm-specific effect to control for this heterogeneity. This model makes strong assumptions: the individual heterogeneity is part of a compound error term and this error term is uncorrelated with the regressors.

As for the participation equation, we assume a normal distribution for the random effects and maximize the likelihood function of our specification including the firm-specific effects (see Guilkey and Murphy, 1993). The integral in the likelihood function can be approximated with the non-adaptive Gauss-Hermite quadrature.[21] The quadrature formula requires that the integrated formula is well approximated by a polynomial. As for the quantity equation, the random effects model is estimated by GLS (dependent and independent variables are transformed using the idiosyncratic and the individual components of the error term).

Table 7 shows the results of the two-part model, including the firm-specific effects, for the total sample and for the subsamples of small and large firms. The random effects probit model is calculated using quadrature. As the panel size increases, the quadrature approximation becomes less accurate.[22] If the results of the estimation change when one changes the number of quadrature points, the results should be dismissed. After checking the magnitude of these changes, we obtain that for most variables, the relative difference between the coefficients using different quadrature points is smaller than 0.01%.[23] So the results of the probit random effects model estimated in this Section can be trusted.

The results for both the participation equation and the quantity equation are similar to those in Table 5 columns (b). The same variables are significant and with the same sign. Although the results are similar to the model without the inclusion of random effects, the tests reject the null hypothesis that the firm-specific effects are zero. For the participation equation, the likelihood-ratio test compares the pool estimator (probit) with the panel estimator. When the panel-level variance component is unimportant, the panel estimator is not significantly different from the pooled estimator. The test rejects the null that the panel-level variance component is equal to zero at 1%. As for the

[21] Further explanation of the estimation method can be found in Greene (1999, Chapter 21). Estimations have been done using the commands `xtprobit` and `xtreg` by the STATA software**.**

[22] We have observations for only two years, so panel size is small and should not present grave quadrature problems.

[23] The software STATA permits checking whether the results of the estimation change when changing the quadrature points by means of the command `quadchk`.

quantity equation, the Breusch and Pagan Lagrange-multiplier test rejects the null hypothesis at 1%. As shown in Table 7, similar results are obtained for the subsamples of small and large firms. According to all we have said until now, we have chosen the two-part model with random effects to carry on the remaining of our analysis.[24]

As for the total sample (first and second columns in Table 7), the firms' size variable is positive and significant in the participation equation indicating the presence of effects associated to large firms even after controlling for the set of possible training determinants. However, it is not significant in the quantity equation.[25]

The percentage of white collars is positive and significant, meaning that firms with more educated workers are more likely to provide training because these workers can take more profit of it. Increasing the percentage of white collars one point increases the firms' expenditure on training per worker by almost 2%. The percentage of white collars is lagged to capture the effect that training is directed to those who have already acquired other knowledge in the past.[26]

In relation to training for technical purposes, firms that use advanced technologies with a medium or high intensity are more likely to provide training: using more complex technology requires more specialized knowledge and, as very specialized skills are not easily found in the labor market, firms may need to provide training.[27] However, we do not find a significant effect of this variable in the quantity equation. The complexity of the adopted technologies is considered contemporaneous to the provision of training. Firms adopt advanced technologies and they intend to incorporate them to the production process as soon as possible. We argue that, when workers need some training to be able to use the new technology, firms will have to provide training not before and not after the adoption of the technologies. If firms provided training before

[24] Given that we have only two years in our analysis, many variables (sector and regional dummies among others) do not show variation over time and then we cannot estimate the fixed effects model. Thus, we cannot apply the Hausmann test to choose whether it is more appropriate to estimate a random effects or a fixed effects model. However, according with the reasoning in Section 5.3, we consider the random effects as possibly an appropriate way to take into account firms' heterogeneity.

[25] The results are in line with Baldwin, Gray and Johnson (1995) and Black and Lynch (1998). Alba-Ramírez (1994) finds a positive effect for the two decisions.

[26] Alba-Ramírez (1994) estimates a probit and tobit models and find a positive and significant effect of this variable. Black and Lynch (1998) estimate a logit model and find that workers' education has a positive and significant impact on computer and teamwork training. For the tobit model, they do not find significant effects.

[27] Baldwin, Gray and Johnson (1995) also find that the probability of training taking place increases as the number of advanced technologies in use increases.

that, workers could leave the firm before it captured the returns from training. If firms provided training after that, the new technology would be idle for a period of time.

Being an innovative firm has a positive and significant effect on the probability of providing training: if firms have to launch a new product or implement a more efficient method of production, the skills of their workers need to be adapted. Changing from being a non-innovative firm to an innovative one increases the expenditure on training per worker 14%. As in the case of the use of advanced technologies, this variable is considered contemporaneous. See Section 5.1, where we discuss different specifications.[28]

The variable on the geographic scope of the market is significant and shows that firms operating in international markets have a higher probability of training their workers than firms operating at national, regional or local markets. However this variable does not seem to have an impact in the decision on the quantity of training.[29]

Being participated by foreign capital also increases the probability of providing training. Increasing the participation of foreign capital in the firm one point increases the firms' expenditure on training per worker around 0.2%. Very often, these firms are multinational firms, more efficient in their management, who employ more qualified workers and who have a more positive attitude toward workers' skills than national firms.[30]

Finally, firms that have a high degree of temporary employment are expected to be less interested in training their workers as they will not be able to capture the returns from it if workers leave their jobs. The variable on the percentage of temporary workers is negative although not significant in the participation equation. However, in the quantity equation, increasing the percentage of temporary workers one point decreases the firms' expenditure on training per worker around 0.6%. Although this is a quite small effect in magnitude, it is significant and with the expected sign.[31] The variable on

[28] Alba-Ramírez (1994) finds a positive and significant effect of this variable defined as in our specification for both the probit and tobit models.

[29] Bartel (1989) estimates a logit model and finds a positive and significant effect of the degree of competition faced by firms. This paper measures the degree of competition through the concentration ratio in the industry, for domestic competence, and through the ratio of net imports (imports minus exports over sales), for foreign competence.

[30] Hughes, O'Connell and Williams (2004) find a positive and significant effect of this variable on the decision of whether to provide training but not on the quantity of training provided.

[31] Black and Lynch (1998) find a negative effect of this variable in the logit model while in the tobit model it is still negative although non-significant. Alba-Ramírez (1994) finds a positive and significant effect of temporary workers under training contracts in the probit model.

the percentage of temporary workers enters the equation without any lag as in Alba-Ramírez (1994).

As for the control variables, the percentage of use of the productive capacity and belonging to a group does not increase the probability to provide more training. Finally, the sets of region and industry dummies are jointly significant.

The fact that firm size is significant in the participation equation, even after controlling for other variables and firm-specific effects, suggests the existence of scale economies in the provision of training as well as other effects associated with firm size. A part from the direct effect of size, the other covariates may have different effects in small and large firms' subsamples, as suggested by the descriptive in Tables 1 and 2. For example, does the increase in the ratio of skilled workers lead to higher probability of training (or more expenditure) in both small and large firms? To further analyze this question we estimate the same equations for the subsamples of small and large firms. Given that small firms are recognized to have more difficulties in accessing training, we are interested in analyzing the impact of these variables in the training decisions and whether they play different roles in firms with different sizes.

Table 7. Estimation the two-part model with random effects for the total sample and the small and large firms' subsamples

	total sample		**small firms' sample**		**large firms' sample**	
	participation eq	quantity eq	participation eq	quantity eq	participation eq	quantity eq
	mg eff	coef	mg eff	coef	mg eff	coef
size	0.6273***	-0.0341	0.6955***	-0.2347**	0.379***	0.0299
	(0.0683)	(0.0484)	(0.1161)	(0.1075)	(0.154)	(0.082)
white	0.0276***	0.0195***	0.0375***	0.0197***	0.0062	0.0177***
	(0.0053)	(0.0038)	(0.0069)	(0.0048)	(0.0082)	(0.0058)
temp	-0.0002	-0.0066*	0.0005	-0.002	-0.0029	-0.0156***
	(0.003)	(0.0034)	(0.0036)	(0.0042)	(0.0062)	(0.0053)
atmed	0.7019***	0.0436	0.8085***	-0.0142	0.4604**	0.1931
	(0.1471)	(0.1123)	(0.1924)	(0.1536)	(0.2367)	(0.1727)
athigh	0.7696***	0.1576	0.8152***	0.1408	0.6219**	0.2403
	(0.2073)	(0.1395)	(0.3249)	(0.2488)	(0.2709)	(0.183)
innov	0.5832***	0.1414*	0.6272***	0.0567	0.529***	0.2244**
	(0.1078)	(0.0795)	(0.1406)	(0.1285)	(0.1701)	(0.0968)
market	0.5413***	0.096	0.397**	0.2674*	0.6401***	0.051
	(0.1282)	(0.0904)	(0.1773)	(0.1543)	(0.1889)	(0.1121)
forgnk	0.0043***	0.0022**	0.0075**	0.0047**	0.0028	0.0013
	(0.0018)	(0.0011)	(0.0032)	(0.0022)	(0.0021)	(0.0013)
prodcap	-0.004	-0.002	-0.0024	-0.0016	-0.0056	-0.0036
	(0.0041)	(0.0031)	(0.0051)	(0.0043)	(0.0075)	(0.0042)
group	0.2078	0.113	0.1339	0.0297	0.0621	0.1983
	(0.1717)	(0.1234)	(0.2435)	(0.193)	(0.2455)	(0.1674)
dyear	-0.1431*	-0.2154***	-0.0656	-0.2004**	-0.3019***	-0.2147***
	(0.0761)	(0.0507)	(0.0973)	(0.0937)	(0.1267)	(0.0587)
cons	-4.0818	4.2892***	-5.0014***	4.3705***	5.7329	4.7874***
	(0.9042)	(0.709)	(1.1196)	(0.9879)	(509.5457)	(0.9618)
H0:dsec=0	yes	yes	yes	yes	yes	yes
	30.81**	39.15**	25.84	40.77***	15.08	47.09***
H0:dreg=0	yes	yes	yes	yes	yes	yes
	41.52***	20.31	36.91***	16.77	13.46	30.35***
H0:RE=0	yes	yes	yes	yes	yes	yes
	230.48***	68.17***	164.98***	15.72***	48.31***	42.65***
num obs	3020	1222	2086	520	934	702
num firms	1538	734	1068	335	493	409
pseudolnL	-1223.5902		-777.5287		-421.5722	

Note: Standard deviation in parentheses
(***) (**) and (*) denote significant at 1%, 5% and 10%.

The results for the participation equation in the case of small firms (third column) are similar to those for the total sample: the same variables are significant and have the expected sign. In the case of large firms (fifth column), only *SIZE*, *ATMED*, *ATHIGH*, *INNOV* and *MARKET* are significant. As for the quantity equation in the case of small firms (fourth column) *SIZE*, *WHITE*, *MARKET* and *FOREIGNK* are significant, while in the case of large firms (sixth column) only *WHITE*, *INNOV* and *TEMP* are. Results suggest the existence of certain differences between small and large firms in their decisions on the quantity of training.

Again, we include the variable on firm size as a regressor. This way, we analyse if, even after separating the two groups and controlling for other covariates, there are differences associated with firm size. This variable has no significant impact on the amount of training for large firms, while it is significant and negative value for small firms. These results suggest the heterogeneity in the training expenditure by size and the necessity for further analysis, as done in Section 6. The differences in the behavior of small and large firms are commented in the following paragraphs.

The percentage of white collars does not determine that large firms decide to provide training, but it does have an impact on the amount of it. This result could be explained by the fact that large firms employ a wide range of employees, and so, *ceteris paribus*, they have a higher probability of providing training to at least one employee. Thus, having more qualified workers does not determine the yes/no decision, but the expenditure per worker.

As for the intensity of use of advanced technologies, these variables have a positive and significant impact for both small and large firms, although the effect in the case of small firms in much larger. Changing from being a non-innovative large firm to an innovative one increases the expenditure on training per worker almost 22%, but in the case of small firms, this variable does not have a significant effect.

Competing in an international market and being participated by foreign capital affect the two decisions in the case of small firms, while in the case of large firms, the former only determines the yes/no decision. This suggests that small firms that operate in international markets or are participated by foreign capital are highly competitive firms that have into special consideration the qualification of their labor force.

Finally, the percentage of temporary workers is only significant and with negative sign in the decision on the quantity of training for large firms. Again, as large firms employ a wide range of workers, it does not affect their probability of providing training but the quantity of it. While in the case of large firms, we observe no significant impact of this variable.

In relation with the control variables on the group and use of the productive capacity, small and large firms do not show differences in behavior. As for the sets of dummy variables on the region and sector, there are differences between small and large firms.

In conclusion, there are certain differences between small and large firms that may explain why large firms provide more training per employee than small ones. In Section 6 we use the Oaxaca-Blinder decomposition to further investigate the underlying reasons of such differences.

6. The Oaxaca-Blinder Decomposition of Training Provision Decisions

6.1. The Oaxaca-Blinder Decomposition in the Two-Part Model

We apply the Oaxaca-Blinder decomposition to analyze whether small and large firms follow different patterns of behavior in their training decisions. It permits decomposing the differences in the yes/no training decision and in the amount of training in two components: differences in the determinants of training and differences in the impact of these determinants. The first component reflects that small and large firms have different characteristics, which are associated to different training levels. The second component reflects the differences in the impact of such characteristics on the training provision by firm size. For example, supposing that small and large firms had the same proportion of qualified workers, would they show a similar propensity to invest in training? This component shows that the origin of the differences in training may arise because of the fact that firms' characteristics may have different impact on their training decisions in small and large firms (i.e. a different coefficient as opposed to different levels in characteristics).

We depart from two auxiliary regressions for small and large firms:

$$\begin{aligned} \hat{T}_S &= F(X'_S \hat{\beta}_S) \\ \hat{T}_L &= F(X'_L \hat{\beta}_L) \end{aligned} \tag{5.7}$$

where T denotes training, both as a discrete (*TR*) or continuous variable (*lnTR*), X is the matrix of the regressors, β is the conforming vector of estimated coefficients and subscripts L and S refer to large firms and small firms respectively.

Notice that $F(\cdot)$ can be both a linear or a non-linear function. A complete decomposition of the two-part model requires decomposing the gap of the variable of interest in the quantity equation, which is a linear model, and the gap of the variable of interest in the participation equations, which is a probit model and so, non-linear. The traditional Oaxaca-Blinder decomposition can only be applied in linear regression models, but it is not possible to perform the detailed decomposition of a nonlinear equation such as the participation equation. In the case of the quantity equation of the two-part model, the standard Oaxaca and Ransom decomposition can be applied. Yun (2004) comments that "the contribution of the differences in characteristics and coefficients of individual variables (i.e. the detailed decomposition) can be easily found when linear equations are used, but not when non-linear equations are used". This author proposes a methodology for the detailed Oaxaca-Blinder decomposition with non-linear functions for each variable.[32]

Yun's methodology consists on finding the contribution of every n-variable to the total difference. The Yun-Oaxaca-Blinder detailed decomposition for non-linear equations is expressed as follows:

$$\hat{T}_L - \hat{T}_S = \sum_{n=1}^{N} W_{\Delta X}^{n}\left[\overline{\Phi(X_L\hat{\beta}_L)} - \overline{\Phi(X_S\hat{\beta}_L)}\right] + \sum_{n=1}^{N} W_{\Delta\beta}^{n}\left[\overline{\Phi(X_S\hat{\beta}_L)} - \overline{\Phi(X_S\hat{\beta}_S)}\right] \tag{5.8}$$

where, in the case of the probit model, Φ is a standard normal cumulative distribution function and $W_{\Delta X}^{n}$ and $W_{\Delta\beta}^{n}$ are the weights for each n-variable.

The key question is finding proper weights for the variables. Yun (2004) suggests evaluating the value of the function using mean characteristics and then using a first order Taylor expansion to linearize Φ around $\overline{X}_L\hat{\beta}_L$ and $\overline{X}_S\hat{\beta}_S$. In this way, he derives the expression for the weights:

$$W_{\Delta X}^{n} = \frac{(\overline{X}_L^{n} - \overline{X}_S^{n})\hat{\beta}_L^{n}}{(\overline{X}_L - \overline{X}_S)\hat{\beta}_L}; \; W_{\Delta\beta}^{n} = \frac{(\hat{\beta}_L^{n} - \hat{\beta}_S^{n})\overline{X}_S^{n}}{(\hat{\beta}_L - \hat{\beta}_S)\overline{X}_S}; \tag{5.9}$$

As we use a variation of the Oaxaca-Blinder decomposition suggested by Oaxaca and Ransom (1994) which does not make any assumption on which is the natural model**,** the decomposition for the participation equation in the two-part model is calculated as follows:

$$\begin{aligned}\hat{\tau}_L - \hat{\tau}_S = {} & \sum_{n=1}^{N} W_{\Delta X}^{n}\left[\overline{\Phi(X_L\hat{\beta}^{*})} - \overline{\Phi(X_S\hat{\beta}^{*})}\right] + \sum_{n=1}^{N} W_{\Delta\beta fav}^{n}\left[\overline{\Phi(X_L\hat{\beta}_L)} - \overline{\Phi(X_L\hat{\beta}^{*})}\right] + \\ & \sum_{n=1}^{N} W_{\Delta\beta disc}^{n}\left[\overline{\Phi(X_S\hat{\beta}^{*})} - \overline{\Phi(X_S\hat{\beta}_S)}\right]\end{aligned} \tag{5.10}$$

[32] As far as we know Yun's detailed decomposition have been only applied so far in a reduced number of labor market studies (Motellón and López-Bazo, 2005; Hernanz and Toharia, 2006).

Linearizing the characteristics and coefficients around $\overline{X}_L \hat{\beta}_L$, $\overline{X}_S \hat{\beta}_S$ and $\overline{X}_S \hat{\beta}^*$, the weights are calculated as:

$$W_{\Delta X}^n = \frac{(\overline{X}_L^n - \overline{X}_S^n)\hat{\beta}^{*n}}{(\overline{X}_L - \overline{X}_S)\hat{\beta}^*} \; ; \; W_{\Delta\beta fav}^n = \frac{(\hat{\beta}_L^n - \hat{\beta}^{*n})\overline{X}_L^n}{(\hat{\beta}_L - \hat{\beta}^*)\overline{X}_L} \; ; \; W_{\Delta\beta disc}^n = \frac{(\hat{\beta}^{*n} - \hat{\beta}_S^n)\overline{X}_S^n}{(\hat{\beta}^* - \hat{\beta}_S)\overline{X}_S} \qquad (5.11)$$

where $\hat{\beta}^*$ is the estimated nondiscriminatory coefficients structure, calculated as a weighted average of the small and large coefficients structure: $\hat{\beta}^* = \Omega\hat{\beta}_L + (I - \Omega)\hat{\beta}_S$, where Ω is specified by: $\Omega = (X'X)^{-1}(X'_L X'_L)$. The subscripts $\Delta\beta fav$ and $\Delta\beta disc$ indicate that the weights correspond to the effect of large firms' advantage and small firms' disadvantage in relation with the non-discriminatory coefficients structure.

The first term at the right hand-side of equation (5.10) reflects training differences due to differences in characteristics. This term is an estimate of the differential in the probability of providing training between small and large firms in the absence of differences in the impact of these characteristics. The second and third terms are estimates of the differential in probability of providing training due to differences in the *impact* of firms' characteristics. Together, they collect the effect of large firms' advantage and small firms' disadvantage in relation with the non-discriminatory coefficients structure. Since we are not particularly interested in distinguishing the advantage and disadvantage effects, but in evaluating the differences in the coefficients as a whole, we will consider these two terms together.

6.2. Results of the Decomposition of the Training Gaps

The main purpose of this paper is analyzing the reasons why small firms provide less training than large firms. The results in Section 5.2 show evidence of certain firm characteristics that determine the probability of providing training and the quantity of resources devoted to this activity. There, we have also shown that the effect of these determinants differs across firms' size. In this Section, we go one step further by trying to assess whether differences in the probability of providing training and differences in the expenditure can be explained just by differences in the level of the determinants of the training provision in small and large firms. Or otherwise, they are in part originated by the different impact of such characteristics on the training provision decisions. To perform such analysis, we apply the detailed decomposition described in Section 6.1.

Tables 8 and 9 show the results for the Oaxaca-Blinder as suggested by Oaxaca and Ransom (1994) for 2001 and 2002. The former, shows the results for the estimation without firm-specific effects and the latter includes firm-specific effects to control for possible heterogeneity among firms. As commented in Section 6.2, the Oaxaca-Blinder decomposition is exact for the OLS estimation and probit model, but it is not for the linear regression and probit models with random effects. Although the decomposition is not exact, we still use the coefficients from the random effects model as it has other strengths (i.e. taking heterogeneity into account). Further research should be done in developing a decomposition that could overcome this limitation of the Oaxaca-Blinder decomposition. Given that some of the determinants of the provision of training are defined as dummy variables, we have applied the Gardeazábal and Ugidos (2004) transformation in the estimation of equations (5.7). In absence of this transformation, it is not possible to distinguish the effects due to the different sets of dummies (see Section 6.2 for further details).

Table 8 shows the main results of the decomposition based on the estimation of the two-part model without firm-specific effects.[33] The differential in the probability of providing training between small and large firms is 0.4 in 2001 and 0.45 in 2002. The decomposition for all the variables together shows that the whole differential can be explained by differences in firms' characteristics. Actually, differences in the impact of characteristics seem to favor small firms. That is, under equal impact of characteristics (i.e. coefficients), the gap in the probability of providing training would have been even larger. However, we are especially interested in the individual decomposition to analyze the contribution of each variable. The proportion of white collars explains a very small part of the differential in the probability of providing training between small and large firms. For this variable, the differences in characteristics, which favor large firms, are compensated by differences in the impact of characteristics, which favor small firms. As for the variables related to technological activities, the use of advanced technologies explains around 11-14% of the differential in the probability of providing training: around 10% is due to differences in characteristics and the remaining is due to differences in the coefficients, both in favor of small firms. This result indicates that small firms make a more intense use of advanced technologies and the coefficients are

[33] Table 8 shows the most relevant results of the decomposition. For more detailed and complete results, see Table A.4 at the Appendix A.3.

also larger. The innovative activity of the firm explains quite a large part of the differential in the probability of providing training: about 15-18% of this differential is due to the fact that large firms innovate more. The binary indicator on whether the firm operates in an international market has also a quite important contribution to explain the probability gap: around 12-13% of it is due to the fact that large firms operate in international markets. The fact that large firms are more participated by foreign capital explains only 6% of the differential. For the last two variables, the differences in the impact of characteristics explain a very small part of the effect in favor of small firms. Also, the proportion of temporary workers explains a very small part of the gap in the probability of providing training.

The differential in the log of expenditure on training per worker between small and large firms is 0.18 in 2001 and 0.28 in 2002. The decomposition for all the variables together shows that in 2001 the whole differential can be explained by differences in firms' characteristics, while differences in the impact of characteristics favor small firms. However, in 2002, differences in characteristics explain around 85% of the differential and the remaining 15% is due to differences in the impact of these characteristics, both effects in favor of large firms. As for the individual decomposition, the proportion of white collars has a very small effect in explaining the differences between small and large firms in their decision on the quantity of training. As for the variables related to technological activities, the use of advanced technologies plays a major role in explaining the differential in the quantity of training: the effect is mainly due to differences in the impact of characteristics in favor of large firms. However, the effect is much larger in magnitude in 2001 (117%) than in 2002 (73%). The innovative activity of the firm also plays an important role in explaining the differences in expenditure on training per worker between small and large firms, both as differences in characteristics and differences in the impact of these characteristics in favor of large firms. The effects are approximately three times as large in 2001 as in 2002. The fact that firms operate in international markets and that they are participated by foreign capital also explain quite an important part of the gap in the quantity of training per worker, both as differences in characteristics and in the impact of these characteristics (the latter with a larger magnitude of the effect in 2001 than in 2002). For the two variables, the portion of the gap explained by differences in characteristics favor of large firms. However, in the case of the market scope of the firm, differences in the impact of characteristics favor large firms, while in the case of the foreign capital participation, differences in the impact of characteristics favor small firms. Although the percentage of temporary workers has a minor contribution in explaining the differential in the probability to provide training, it plays a major role in explaining the differential in the quantity of training and it is mainly due to differences in the impact of characteristics in favor of small firms (111% in 2001 and 61% in 2002). In other words, if small and large firms had the same proportion of temporary workers, *ceteris paribus*, the gap in the probability of providing training between small and large firms would be even wider.

Table 8. Decomposition for the two-part model. Estimation without firm-specific effects

	2001				**2002**			
	participation eq		quantity eq		participation eq		quantity eq	
Training differential	**0.40143463**		**0.18120887**		**0.45014563**		**0.28994218**	
	Charact	Impact	Charact	Impact	Charact	Impact	Charact	Impact
Total	0.461	-0.059	0.25	-0.069	0.461	-0.011	0.248	0.041
	114.87%	-14.87%	138.14%	-38.14%	102.46%	-2.46%	85.55%	14.45%
White collars	0.018	-0.021	-0.004	-0.004	0.021	-0.006	0.013	-0.004
	4.57%	-5.30%	-2.40%	-1.97%	4.64%	-1.48%	4.55%	-1.31%
Advanced Technology	-0.04	-0.017	0.002	0.213	-0.041	-0.01	0.005	0.212
	-9.93%	-4.07%	0.87%	117.52%	-9.06%	-2.12%	1.76%	72.92%
Innovation	0.071	0.001	0.058	0.045	0.065	0.003	0.031	0.026
	17.68%	0.17%	31.90%	24.92%	14.53%	0.73%	10.81%	8.80%
International Market	0.051	-0.009	0.034	0.036	0.054	-0.009	0.035	0.034
	12.73%	-2.26%	18.98%	20.18%	11.95%	-1.97%	12.23%	11.63%
Foreign capital	0.025	-0.006	0.054	-0.109	0.027	-0.002	0.064	-0.101
	6.16%	-1.44%	29.77%	-60.35%	5.95%	-0.59%	22.24%	-34.69%
Temporary workers	0.001	-0.007	0.005	-0.202	0.001	-0.003	0.021	-0.177
	0.19%	-1.76%	2.73%	-111.38%	0.19%	-0.70%	7.12%	-61.12%

In Table 9 we offer the main results of the decomposition based on the estimation of the two-part model including firm-specific effects.[34] It is important to notice that the results under pooled and panel data are most similar in the decomposition for the participation equation, while they show slight differences in the individual decomposition for the quantity equation.

The differential in the probability of providing training between small and large firms is 0.5 in 2001 and 0.54 in 2002. The decomposition for all the variables together shows that differences in characteristics explain the whole differential and that

[34] Table 9 shows the most relevant results of the decomposition. For more detailed and complete results, see Table A.5 at the Appendix A.3.

differences in the impact of characteristics favor small firms. As before, the proportion of white collars explains a very small part of the differential in the probability of firms' providing training. The use of advanced technologies explains around 16-19% of the gap: around 11% is due to differences in characteristics and the remaining is due to differences in the coefficients, both effect favoring again small firms. And the fact that large firms innovate more explains 11-13% of the differential. Around 13-14% of the gap in the probability of providing training it is due to the fact that large firms operate in international markets more than small ones. The higher participation of foreign capital in large firms explains only 6% of the differential and the percentage of temporary workers explains an almost negligible part of the gap.

As for the quantity equation, the differential in the log of expenditure on training per worker between small and large firms is 0.22 in 2001 and 0.28 in 2002. Differences in characteristics explain the whole differential and differences in the impact of characteristics favor small firms in 2001. While in 2002, differences in characteristics explain around 79% of the differential and 21% is due to differences in the impact of characteristics, both effects in favor of large firms. Around 11-13% of the differential in the quantity of training is due to differences in the impact of characteristics in favor of small firms in the percentage of white collars (i.e. if small and large firms had the same proportion of white collars, *ceteris paribus*, the gap would be even wider). This result differs from that obtained in Table 8, where this variable had a very small effect in explaining the gap. Differences in the impact of characteristics in favor of large firms of the variable on the use of advanced technologies play an important role in explaining the differential in the quantity of training per worker: 92% in 2001 and 74% in 2002. In contrast with the results in Table 8, in 2001, this variable does not explain the whole differential. The innovative activity of the firm explains a relevant portion of the gap, both as differences in characteristics and differences in the impact of characteristics in favor of large firms, and the magnitude of these effects is approximately twice as large in 2001 as in 2002. The international scope of the market and the participation of foreign capital explain an important part of the differential in the amount of training per worker. As in the case without firm-specific effects, the gap is explained by differences in characteristics and differences in the impact of characteristics of these variables. As before, the percentage of temporary workers plays an important role in explaining the differential and it is mainly due to differences in the impact of characteristics in favor of small firms (101% in 2001 and 72% in 2002).

Table 9. Decomposition for the two-part model. Estimation including firm-specific effects

	2001				2002			
	participation eq		quantity eq		participation eq		quantity eq	
Training differential	**0.49120915**		**0.22525114**		**0.54167439**		**0.27912344**	
	Charact	Impact	Charact	Impact	Charact	Impact	Charact	Impact
Total	0.571	-0.079	0.238	-0.012	0.578	-0.036	0.22	0.059
	116.26%	-16.26%	105.47%	-5.47%	106.71%	-6.71%	78.97%	21.03%
White collars	0.026	-0.03	-0.004	-0.029	0.03	-0.023	0.013	-0.032
	5.21%	-6.21%	-1.91%	-12.74%	5.47%	-4.16%	4.67%	-11.25%
Advanced Technology	-0.057	-0.035	-0.009	0.208	-0.06	-0.028	-0.005	0.206
	-11.68%	-7.16%	-4.13%	92.31%	-11.02%	-5.12%	-1.84%	73.82%
Innovation	0.067	0	0.042	0.061	0.063	0.003	0.023	0.035
	13.70%	0.08%	18.76%	27.06%	11.64%	0.45%	8.21%	12.30%
International Market	0.068	-0.014	0.042	0.036	0.073	-0.014	0.043	0.033
	13.87%	-2.80%	18.50%	15.82%	13.46%	-2.62%	15.39%	11.80%
Foreign capital	0.028	-0.008	0.052	-0.1	0.031	-0.006	0.063	-0.092
	5.73%	-1.59%	23.28%	-44.45%	5.72%	-1.05%	22.46%	-33.14%
Temporary workers	0	-0.005	0.005	-0.228	0	-0.004	0.023	-0.201
	0.05%	-0.99%	2.43%	-101.39%	0.05%	-0.61%	8.18%	-72.18%

To summarize, the proportion of white collars explains a very small part of the differential in the probability of firms' providing training. But it explains a wider part of the gap of the quantity of training as differences in the impact of characteristics between small and large firms. In relation with technological activities, the use of advanced technologies and being an innovative firm explain a bit more than 10% of the differential in the probability of providing training. The former explains a large part of the differential in the quantity of training differences in the impact of characteristics in favor of large firms. While the latter explains a relevant portion of this gap, both as differences in characteristics and differences in the impact of characteristics also in favor of large firms. As for the market scope of the firm and the participation of foreign capital, differences in characteristics explain a modest part of the differential in the probability of providing training. However, these variables explain a larger part of the gap in the quantity of training per worker, both as by differences in characteristics and differences in the impact of characteristics. Finally, the percentage of temporary workers plays an important role in explaining the differential and it is mainly due to differences in the impact of characteristics in favor of small firms.

7. Conclusions

In this paper we try to assess the reasons why small firms provide less training than their larger counterparts. Our hypothesis is that large firms are associated to certain characteristics (or determinants) which make them more likely to provide training and spend more on training per worker.

First, we presented theoretical arguments and previous empirical evidence supporting the hypothesis that training is associated to certain characteristics such as the previous qualification of the labor force, the technological complexity of the productive process, the innovative capacity of the firm, the fact that firms operate in an international market, the participation of foreign capital in the firm and the percentage of temporary workers. We also presented evidence in favor of the fact that large firms invest more on training and they are more associated to these characteristics than small firms.

Next, we estimate the heckit model, which encompasses the two-part model, to analyze if these characteristics explain the decision on whether to provide training or not and how much to spend on it. After a discussion on the two models from a theoretical and applied perspective, we consider that the two-part model seems more appropriate to model firms' decisions on training. Then, we perform the remaining of our analysis on the basis of the two-part model. Departing from the idea that small and large firms follow different patterns in their training decisions, we estimate the two subsamples separately. We obtain that the percentage of white collars increases the probability of providing training in the case of small firms, but not in large firms. The intensity of use of advanced technology has a positive and significant impact for both small and large firms, although the effect in the case of small firms in much larger. Being a large innovative firm has a positive effect on the quantity of training, but it is not the case of small firms. Competing in an international market and being participated by foreign capital affect the two decisions in the case of small firms. While in the case of large firms, only the former determines the yes/no decision and the latter has no effect. The percentage of temporary workers increases the expenditure on training per worker in the case of large firms, but not in small firms.

Finally, the Oaxaca-Blinder decomposition assesses the relative contribution of these determinants to explain the different patterns of training provision between small and large firms. These differences are in part associated to the above-mentioned determinants, especially those related with the technological activity, the degree of competition and the participation of foreign capital. Concretely, the fact that large firms make a more intense use of advanced technologies and innovate more explains a large part of the differential in the decision on the provision of training and the quantity of it. The market scope of the firms, reflecting the degree of competition that they face, and the participation of foreign capital also explain an important part of the differential in the quantity of training per worker.

All in all, we obtain that small firms face more restrictions in their access to training. The technological activity, the degree of competition and the participation of foreign capital are the main reasons explaining the fact that small firms provide less training. Our results suggest that the training provision differential is associated with the firms' need to update the skills of their employees so that they acquire specific knowledge to use the new technologies and also to make the firm more competitive in a foreign environment. This can be seen as a limitation for small firms to become more competitive: they have a more restricted access to training so as to take more profit from their technological activity and their effort to compete in a foreign market. Or, in other words, small firms have a more limited access to a tool that permits adapting the skills of their employees for becoming more competitive.

References

Alba-Ramirez, Alfonso, (1994). "Formal Training, Temporary Contracts, Productivity and Wages in Spain," Oxford Bulletin of Economics and Statistics, Blackwell Publishing, vol. 56(2), pages 151-70, May.

Almeida, R. and Carneiro, P. (2005), "On-the-Job Training: Estimating Costs and Returns Using Firm Level Data", mimeo.

Barrett, A and O'Connell, PJ (2001), “Does Training Generally Work? The Returns to In-Company Training”. Industrial and Labor Relations Review, 54(3), pp 647-662.

Barry, F., Görg, H. and Strobl, E. (2004). "Multinationals and Training: Some Evidence from Irish Manufacturing Industries," Scottish Journal of Political Economy, Scottish Economic Society, vol. 51(1), pages 49-61, 02.

Bartel, A. (1989), “Formal employee training programs and their impact on labor productivity: evidence from a human resources survey”, NBER Working Paper Series, No. 3026.

Bartel, A. (1994), "Productivity gains from the implementation of employee training programs” Industrial Relations, Volume 33, Number 4, October 1994

Black, S.E. and L.M. Lynch (1996), “Human-capital investments and productivity”, American Economic Review, Papers and Proceedings, May.

Black, S.E. and L.M. Lynch (1998), "Beyond the Incidence of Training: Evidence from a National Employers' Survey", Industrial and Labor Relations Review, Oct. 1998, pp. 64-81.

Black, D., B. Noel, and Z. Wang (1999). “On-the-Job Training, Establishment Size, and Firm Size: Evidence for Economies of Scale in the Provision of Human Capital,” Southern Economic Journal, July 66(1) 82-100

Blundell, R., Dearden, L., Meghir, C. and Sianesi, B. (1999). Human Capital Investment: The Returns from Education and Training to the Individual, the Firm and the Economy. Fiscal Studies, vol. 20, no. 1, pp. 1–23

Busom, I. (2000), “An empirical evaluation of the effects of R&D subsidies,” Economic Innovation and New Technology, Vol 9, 111-148.

Cameron, A.C. and P.K. Trivedi (2005), Microeconometrics, Princeton University Press, Princeton.

Cragg, J. G. (1971). “Some Statistical Models for Limited Dependent Variables with Application to the Demand for Durable Goods.” Econometrica 39: 829-44.

Crespo, J. and Sanz, I. (2000): "La formación continua en España: implicaciones de política económica. Papeles de Economía Española 86. Págs. 280-294.

David, P. A., Hall, B.H. and Toole, A. A., (2000), "Is public R&D a complement or substitute for private R&D? A review of the econometric evidence," Research Policy, Elsevier, vol. 29(4-5), pages 497-529, April.

Dearden, L., Reed, H. and Van Reenen, J. (2000), "Who Gains when Workers Train? Training and Corporate Productivity in a Panel of British Industries", IFS Working Paper No. W00/04, March.

Duan, N., Manning, W. G., Morris, C. and Newhouse, J. (1983), "A Comparison of Alternative Models for the Demand of Medical Care", Journal of Business and Economic Statistics 1:115-26.

Heckman, James J. (1979), "Sample Selection Bias as a Specification Error", Econometrica, Vol. 47, No. 1 (Jan., 1979) , pp. 153-162

Holtmann, Alphonse G. and Todd L. Idson (1991), "Employer Size and On-the-Job Training Decisions", Southern Economic Journal, Vol. 58, No. 2 (Oct., 1991) , pp. 339-355

Holzer, H.J., R.N. Block, M. Cheatham and J.H. Knott (1993), Are training subsidies for firms effective? The Michigan experience, Industrial and Labor Relations Review, 46, pp. 625-636.

Lach, S. (2002), "Do R&D Subsidies Stimulate or Displace Private R&D? Evidence from Israel," Journal of Industrial Economics, Blackwell Publishing, vol. 50(4), pages 369-90, December.

Leung, S. F., and S. Yu, "On the choice between sample selection and two-part models", Journal of Econometrics, 72, 1996:197-229.

Leuven, E. and Oosterbeek, H. (2004), "Evaluating the Effect of Tax Deductions on Training," Journal of Labor Economics, University of Chicago Press, vol. 22(2), pages 461-488, April.

Lisbon Strategy, "Convergencia y Empleo. Programa Nacional de Reformas para España", Octubre (2005), http://www.fundaciontripartita.org

Motellón, E. and López-Bazo, E. (2005), "Discriminación por género en el acceso a la contratación indefinida", VIII Encuentro de Economía Aplicada, Murcia.

Otero, C. Muñoz, A. and Fernández-Rodríguez, C.J. (2002), "Vocational Education and Training in Spain", Cedefop, reference series 38.

Parellada M., Sáez F., Sanromà E., Torres C. (1999), La formación continua en las empresas españolas y el papel de las universidades. Biblioteca Civitas Economía y Empresa: colección economía.

Peraita, C. (2000). Características de la formación en la empresa española. Papeles de Economía Española. 80, pp. 295-307.

Planas, J. (2005), "Vocational training in Spain: Changes in the model of skill production and in management modalities", Berufs-Wirtsch.pädagog. issue 7 of www.bpwat.de

Planas, J. and Rifa, J. (2003), "Consolidación y desarrollo de la formación continua en España". Madrid: Fundación Tripartita.

Appendix

A.1. Description of the Variables

- Training is measured as a discrete variable (*dTR*), according to whether the firm provides continuous training, and as a continuous variable, that is, the log of the real expenditure on continuous training per worker (*lnTR*). Continuous training is measured as the external expenses on training per worker, including fives different types of training: computation and information technologies, foreign languages, sales and marketing, engineering and technical training and other issues (and expressed in 2001 real euros).
- *SIZE* is defined as the total number of employees and measured as the number of full time employees plus the number of part time employees divided by two (both on December 31st) plus the number of temporary employees.
- *WHITE* is the percentage of white collars in the firm including those employees with bachelor or a higher level of studies. Data on white collars is not available in 2000 and 2001 as it is not assumed to change substantially every year. We interpolate the proportion of white collars, making the assumption that they increase or decrease linearly. For the firms that entered the survey in 2000 and 2001, we use data on this year. For the firms that entered the survey in previous years, we interpolate the percentage of white collars in 2000 and 2001, using the corresponding values for every firm in 1998 and 2002 and making the assumption that they increase or decrease linearly.
- The intensity of use of advanced technologies is measured by a set of three dummy variables specified as *ATLOW*, *ATMED* or *ATHIGH*, when firms use 0-1, 2-3 or 4-5 advanced technologies respectively. The survey has questions on whether the following technologies are used by the firm: Computer Numerically Controlled (CNC) machines and tools, Robots, Computer-aided design (CAD), Combination of the previous systems by central computer (CAM, flexible manufacturing systems, etc) and Local Area Network (LAN) for factory use. In the dataset, these data is only available every four years, as it is not supposed to change yearly, and so, we assumed to be constant between 2001 and 2002.
- *INNOV* is a dummy variable that takes value one if the firm has introduced a product or a process innovation (*PRODUCT*, *PROCESS*).

- *MARKET* is defined as a dummy on the geographical scope of the firms' main market. It takes values 1 when the firm operates in an international market. And it takes values zero, when it is local, province, regional or national.
- *FOREIGNK* is the percentage of foreign-owned capital of the firm.
- *TEMP* is defined as the percentage of temporary workers in the firm at the end of 2001 and 2002. When the firm reports that the number of temporary employees has changed considerably, it is computed as the average of temporary employees at the end of every quarter.
- *PRODCAP* is the percentage of the productive capacity used by the firm and it is a question directly asked in the survey.
- *GROUP* is a dummy on whether the firm belongs to a group of firms.
- *DSEC* is a set of 20 dummy variables according to the National Classification of Economic Activities (NACE93). The excluded category is "Office machines, computer equipments, process equipments, optics and similar".
- *DREG* is a set of 17 dummy variables by CCAA. The omitted category is "La Rioja". Due to lack of variability, we consider all the firms situated in the "Balearic Islands" and "Canary Islands" as a single category.
- *DYEAR* is a dummy variable that takes value 1 for firms in 2001 and 0 in 2002.

A.2. Estimation of the Two-Part and Heckit models. Complementary results

Table A.1. Estimation of the two-part and heckit models. Lagged product innovations and contemporaneous process innovations

	(a)					**(b)**				
	Product (t-1) // Process (t)					**Innov (t-1 // t)**				
	Participation eq		Quantity eq			Participation eq		Quantity eq		
			BSSM		two-part model			BSSM		two-part model
	mg eff	coef	mg eff	coef	coef	mg eff	coef	mg eff	coef	coef
size	0.1265***	0.3357***	0.5028	0.0409	-0.0285	0.1302***	0.3464***	0.5123	0.0648	-0.0318
	(0.0108)	(0.0289)		(0.0696)	(0.0396)	(0.0107)	(0.0286)		(0.0747)	(0.0394)
white	0.0053***	0.014***	0.0285	0.0223***	0.0197***	0.0052***	0.0139***	0.0279	0.0227***	0.0193***
	(0.0009)	(0.0024)		(0.0038)	(0.0033)	(0.0009)	(0.0024)		(0.0039)	(0.0033)
temp	-0.0002	-0.0004	-0.0029	-0.0061***	-0.0057*	-0.0002	-0.0005	-0.0031	-0.0065***	-0.0059**
	(0.0006)	(0.0015)		(0.0024)	(0.003)	(0.0006)	(0.0015)		(0.0025)	(0.003)
atmed	0.1302***	0.3402***	0.5267	0.0664	-0.0155	0.1336***	0.3495***	0.5398	0.1036	-0.0096
	(0.0251)	(0.0652)		(0.1122)	(0.09)	(0.025)	(0.0649)		(0.1163)	(0.0896)
athigh	0.1457***	0.375***	0.6372	0.1761	0.1029	0.1504***	0.3874***	0.6528	0.2104*	0.1085
	(0.0368)	(0.0932)		(0.1227)	(0.1094)	(0.0367)	(0.0929)		(0.1266)	(0.1094)
product	0.1099***	0.2869***	0.4924	0.1698*	0.1189					
	(0.0259)	(0.0668)		(0.0894)	(0.0787)	0.1694***	0.4501***	0.7823	0.3855***	0.2614***
process	0.1553***	0.4057***	0.6875	0.2186**	0.1438*	(0.0219)	(0.0586)		(0.1149)	(0.0821)
	(0.0239)	(0.0621)		(0.0993)	(0.0783)					
market	0.1116***	0.2924***	0.4874	0.1456	0.0811	0.1128***	0.2961***	0.4876	0.1661*	0.0788
	(0.0245)	(0.0636)		(0.0958)	(0.0819)	(0.0244)	(0.0634)		(0.0985)	(0.0819)
foreignk	0.0011***	0.0029***	0.0052	0.0028***	0.0023**	0.001***	0.0028***	0.005	0.003***	0.0024***
	(0.0003)	(0.0009)		(0.001)	(0.0009)	(0.0003)	(0.0009)		(0.001)	(0.0009)
prodcap	-0.001	-0.0028	-0.0042	-0.0005	0.0001	-0.0011	-0.0028	-0.0043	-0.0007	0
	(0.0008)	(0.0021)		(0.003)	(0.0029)	(0.0008)	(0.0021)		(0.003)	(0.0029)
group	0.0385	0.1017	0.2049	0.1501	0.1214	0.0366	0.0969	0.1998	0.1674*	0.1292
	(0.0301)	(0.0791)		(0.1017)	(0.099)	(0.0299)	(0.079)		(0.1028)	(0.099)
dyear	-0.0403**	-0.1071**	-0.2582	-0.2791***	-0.2622***	-0.0416**	-0.1108**	-0.264	-0.295***	-0.2699***
	(0.021)	(0.0559)		(0.0732)	(0.0723)	(0.0209)	(0.0558)		(0.0743)	(0.0719)
cons		-2.033***		3.2284***	3.955***		-2.0382***		2.8695***	3.9481***
		(0.4114)		(0.856)	(0.6089)		(0.4101)		(0.8945)	(0.605)
H0:dsec=0	yes		yes		yes	yes		yes		yes
	46.68***		55.84***		3.05***	47.63***		56.86***		3.16***
H0:dreg=0	yes		yes		yes	yes		yes		yes
	65.00***		29.66***		1.84**	67.33***		30.35***		1.83**
num obs	3020		1222		1222	3020		1222		1222
Pseudo R	0.3472		-		0.1796	0.3436				0.1821
pseudolnL	-1330.4748		-		-	-1337.7413				-
rho	-		0.3115		-	-		0.4126		-
sigma2	-		1.2549		-	-		1.2775		-
sigma12	-		0.3909		-	-		0.5271		-
			(0.3301)					(0.3506)		

Note: Standard deviation in parentheses
(***) (**) and (*) denote significant at 1%, 5% and 10%.

Table A.2. Estimation the heckit model and the two-part model for the total sample and the small and large firms' subsamples

	total sample					**small firms' sample**					**large firms' sample**				
	Participation eq		Quantity eq			Participation eq		Quantity eq			Participation eq		Quantity eq		
			BSSM		two-part model			BSSM		two-part model			BSSM		two-part model
	mg eff	coef	mg eff	coef	coef	mg eff	coef	mg eff	coef	coef	mg eff	coef	mg eff	coef	coef
size	0.1318***	0.3507***	0.5213	0.0812	-0.0246	0.0982***	0.3657***	0.354	0.0225	-0.2112**	0.0704***	0.2439***	0.346	-0.1194	0.009
	(0.0107)	(0.0285)		(0.0752)	(0.0395)	(0.0127)	(0.0479)		(0.2222)	(0.098)	(0.0242)	(0.0845)		(0.1233)	(0.064)
white	0.0053***	0.0142***	0.0285	0.0236***	0.0198***	0.0051***	0.019***	0.0239	0.0307***	0.0199***	0.0003	0.001	0.0168	0.019***	0.0197***
	(0.0009)	(0.0024)		(0.004)	(0.0033)	(0.0008)	(0.0029)		(0.0106)	(0.0045)	(0.0013)	(0.0046)		(0.0051)	(0.005)
atmed	0.1324***	0.3463***	0.5373	0.1167	-0.0043	0.1136***	0.3911***	0.4377	0.1817	-0.0648	0.0687**	0.2431**	0.4065	-0.0317	0.1277
	(0.025)	(0.065)		(0.1158)	(0.0898)	(0.0248)	(0.0799)		(0.2546)	(0.1353)	(0.0346)	(0.1251)		(0.1806)	(0.1283)
athigh	0.1511***	0.3893***	0.6592	0.2274*	0.1164	0.1239***	0.4019***	0.5157	0.3258	0.0932	0.1013***	0.3715***	0.6347	-0.0089	0.216
	(0.0366)	(0.0927)		(0.1271)	(0.1094)	(0.0476)	(0.1382)		(0.2956)	(0.2077)	(0.0368)	(0.1439)		(0.2197)	(0.1374)
innov	0.1668***	0.4411***	0.7456	0.3209***	0.1935***	0.14***	0.4855***	0.584	0.4065	0.115	0.1171***	0.3934***	0.729	0.0076	0.2484***
	(0.022)	(0.0585)		(0.1104)	(0.0806)	(0.0225)	(0.0738)		(0.2837)	(0.1287)	(0.0315)	(0.1035)		(0.2047)	(0.1031)
market	0.1109***	0.2912***	0.48	0.1719*	0.0792	0.0613***	0.217***	0.3002	0.3851**	0.2569*	0.1055***	0.3612***	0.4993	-0.2017	0.0306
	(0.0244)	(0.0634)		(0.098)	(0.0819)	(0.0251)	(0.0846)		(0.1796)	(0.1443)	(0.0307)	(0.1048)		(0.1987)	(0.1044)
forgnk	0.001***	0.0027***	0.0049	0.0029***	0.0023***	0.0012***	0.0046***	0.0058	0.0073***	0.0051***	0.0005	0.0019	0.0035	0.0001	0.0013
	(0.0003)	(0.0009)		(0.001)	(0.0009)	(0.0004)	(0.0014)		(0.0028)	(0.0019)	(0.0003)	(0.0012)		(0.0015)	(0.001)
temp	-0.0002	-0.0005	-0.0032	-0.0065***	-0.006**	0.0001	0.0003	-0.0002	-0.0025	-0.0022	-0.001	-0.0033	-0.0153	-0.0118***	-0.0143***
	(0.0006)	(0.0015)		(0.0025)	(0.0031)	(0.0005)	(0.0017)		(0.0035)	(0.0039)	(0.001)	(0.0035)		(0.0046)	(0.0047)
prodcap	-0.0012	-0.0032	-0.0049	-0.0009	0	-0.0008	-0.0029	-0.0033	-0.0027	-0.0008	-0.0011	-0.0039	-0.0062	0.0009	-0.001
	(0.0008)	(0.0021)		(0.003)	(0.0029)	(0.0007)	(0.0025)		(0.0047)	(0.004)	(0.0013)	(0.0044)		(0.005)	(0.0041)
group	0.0361	0.0958	0.1955	0.1624	0.1218	0.0144	0.053	0.0568	0.0285	-0.0023	0.01	0.0343	0.2201	0.2012	0.2172*
	(0.0299)	(0.079)		(0.1032)	(0.0991)	(0.0295)	(0.1067)		(0.1736)	(0.1666)	(0.0388)	(0.1324)		(0.1509)	(0.1305)
dyear	-0.0307	-0.0816	-0.2142	-0.2731***	-0.2525***	-0.0077	-0.0286	-0.0704	-0.2305**	-0.2122	-0.056**	-0.1945**	-0.4706	-0.1525	-0.2626***
	(0.0208)	(0.0555)		(0.0735)	(0.0719)	(0.0185)	(0.0691)		(0.1201)	(0.115)	(0.0278)	(0.097)		(0.131)	(0.09)
cons		-2.0382***		2.8695***	3.9481***		-2.0382***		2.8695***	3.9481***		4.4786		5.3329***	4.5951***
		(0.4101)		(0.8945)	(0.6050)		(0.4101)		(0.8945)	(0.605)		(4.567)		(1.2953)	(0.7253)
H0:dsec=0	Yes		Yes		Yes	yes		yes		yes	yes		yes		yes
	47.80***		56.89***		3.11***	39.54***		31.80**		2.33***	24.26		39.99***		4.95***
H0:dreg=0	yes		yes		yes	yes		yes		yes	yes		yes		yes

	67.08***	30.48***	1.83**	64.91***	18.82	1.55*	58.81***	20.15	2.93***
num obs	3020	1222	1222	2086	520	520	934	702	702
Pseudo R	0.3431	-	0.1789	0.2658	-	0.1902	0.1487	-	0.2674
pseudolnL	-1338.8303	-	-	-860.0167	-	-	-445.7285	-	-
rho	-	0.4425	-	-	0.6575	-	-	-1.0000	-
sigma2	-	1.2888	-	-	1.4597	-	-	1.4874	-
sigma12	-	0.5703	-	-	0.9597	-	-	-1.4874	-
		(0.3484)			(0.8239)			(1.0214)	

Note: Standard deviation in parentheses
(***) (**) and (*) denote significant at 1%, 5% and 10%.

Table A.3. Empirical mean squared error test to choose between the heckit and the two-part model

	total sample						**small firms' sample**						**large firms' sample**					
	2PM is “true” model			**Heckit is “true” model**			**2PM is “true” model**			**Heckit is “true” model**			**2PM is “true” model**			**Heckit is “true” model**		
	mse 2PM	mse Heckit	choice	mse 2PM	mse Heckit	choice	mse 2PM	mse Heckit	choice	mse 2PM	mse Heckit	choice	mse 2PM	mse Heckit	choice	mse 2PM	mse Heckit	choice
size	1.56E-03	1.68E-02	2pm	1.79E-03	5.65E-03	2pm	9.61E-03	1.04E-01	2pm	1.85E-02	4.94E-02	2pm	4.10E-03	3.17E-02	2pm	4.86E-03	1.52E-02	2pm
white	1.09E-05	3.04E-05	2pm	1.09E-05	1.58E-05	2pm	1.98E-05	2.29E-04	2pm	1.99E-05	1.13E-04	2pm	2.50E-05	2.61E-05	2pm	2.50E-05	2.56E-05	2pm
atmed	8.06E-03	2.80E-02	2pm	8.46E-03	1.34E-02	2pm	1.83E-02	1.26E-01	2pm	2.98E-02	6.48E-02	2pm	1.65E-02	5.80E-02	2pm	1.82E-02	3.26E-02	2pm
athigh	1.20E-02	2.85E-02	2pm	1.22E-02	1.62E-02	2pm	4.31E-02	1.41E-01	2pm	5.28E-02	8.74E-02	2pm	1.89E-02	9.89E-02	2pm	2.53E-02	4.83E-02	2pm
innov	6.49E-03	2.84E-02	2pm	6.97E-03	1.22E-02	2pm	1.66E-02	1.65E-01	2pm	3.87E-02	8.05E-02	2pm	1.06E-02	9.98E-02	2pm	1.86E-02	4.19E-02	2pm
market	6.71E-03	1.82E-02	2pm	6.84E-03	9.61E-03	2pm	2.08E-02	4.87E-02	2pm	2.16E-02	3.23E-02	2pm	1.09E-02	9.34E-02	2pm	1.77E-02	3.95E-02	2pm
foreignk	8.46E-07	1.53E-06	2pm	8.46E-07	1.08E-06	2pm	3.63E-06	1.28E-05	2pm	3.63E-06	7.86E-06	2pm	1.09E-06	3.68E-06	2pm	1.09E-06	2.31E-06	2pm
temp	9.31E-06	6.42E-06	h	9.31E-06	6.06E-06	h	1.52E-05	1.25E-05	h	1.52E-05	1.24E-05	h	2.25E-05	2.75E-05	2pm	2.25E-05	2.15E-05	h
Controls																		
prodcap	8.43E-06	9.81E-06	2pm	8.43E-06	9.02E-06	2pm	1.58E-05	2.56E-05	2pm	1.58E-05	2.21E-05	2pm	1.65E-05	2.93E-05	2pm	1.65E-05	2.54E-05	2pm
group	9.82E-03	1.23E-02	2pm	9.82E-03	1.06E-02	2pm	2.78E-02	3.11E-02	2pm	2.78E-02	3.01E-02	2pm	1.70E-02	2.30E-02	2pm	1.71E-02	2.28E-02	2pm
dsec1	1.44E-01	2.13E-01	2pm	1.49E-01	1.83E-01	2pm	3.31E-01	6.68E-01	2pm	4.44E-01	5.18E-01	2pm	2.65E-01	4.51E-01	2pm	3.00E-01	4.49E-01	2pm
dsec2	9.71E-02	1.37E-01	2pm	9.87E-02	1.36E-01	2pm	2.17E-01	3.36E-01	2pm	2.31E-01	3.30E-01	2pm	1.44E-01	4.13E-01	2pm	2.16E-01	3.76E-01	2pm
dsec3	1.44E-01	2.27E-01	2pm	1.51E-01	2.00E-01	2pm	4.44E-01	9.26E-01	2pm	6.77E-01	7.07E-01	2pm	1.66E-01	4.69E-01	2pm	2.58E-01	4.69E-01	2pm
dsec4	9.88E-02	1.38E-01	2pm	1.00E-01	1.38E-01	2pm	2.22E-01	3.04E-01	2pm	2.29E-01	3.04E-01	2pm	1.63E-01	5.39E-01	2pm	3.04E-01	4.35E-01	2pm
dsec5	1.35E-01	3.01E-01	2pm	1.62E-01	2.95E-01	2pm	2.38E-01	5.09E-01	2pm	3.11E-01	5.08E-01	2pm	1.59E-01	1.33E+00	2pm	1.53E+00	1.24E+00	h
dsec6	1.59E-01	1.98E-01	2pm	1.61E-01	1.98E-01	2pm	2.86E-01	4.30E-01	2pm	3.07E-01	4.28E-01	2pm	2.36E-01	7.48E-01	2pm	4.98E-01	5.88E-01	2pm
dsec7	1.05E-01	1.57E-01	2pm	1.08E-01	1.53E-01	2pm	2.71E-01	3.86E-01	2pm	2.84E-01	3.73E-01	2pm	1.46E-01	4.19E-01	2pm	2.21E-01	3.98E-01	2pm
dsec8	1.04E-01	1.52E-01	2pm	1.06E-01	1.46E-01	2pm	2.58E-01	3.46E-01	2pm	2.66E-01	3.45E-01	2pm	1.43E-01	4.01E-01	2pm	2.09E-01	3.83E-01	2pm
dsec9	8.32E-02	1.50E-01	2pm	8.76E-02	1.29E-01	2pm	2.03E-01	4.04E-01	2pm	2.43E-01	3.31E-01	2pm	1.23E-01	3.69E-01	2pm	1.83E-01	3.43E-01	2pm
dsec10	9.31E-02	1.66E-01	2pm	9.84E-02	1.43E-01	2pm	2.01E-01	3.98E-01	2pm	2.40E-01	3.41E-01	2pm	1.46E-01	4.26E-01	2pm	2.25E-01	3.92E-01	2pm
dsec11	1.25E-01	1.48E-01	2pm	1.25E-01	1.48E-01	2pm	4.72E-01	3.85E-01	h	4.79E-01	3.82E-01	h	1.71E-01	4.62E-01	2pm	2.56E-01	4.04E-01	2pm
dsec12	1.01E-01	1.84E-01	2pm	1.08E-01	1.53E-01	2pm	2.81E-01	5.85E-01	2pm	3.73E-01	4.52E-01	2pm	1.50E-01	3.84E-01	2pm	2.05E-01	3.76E-01	2pm
dsec13	9.17E-02	1.59E-01	2pm	9.63E-02	1.36E-01	2pm	2.02E-01	4.36E-01	2pm	2.57E-01	3.45E-01	2pm	1.50E-01	3.68E-01	2pm	1.98E-01	3.64E-01	2pm
dsec14	8.61E-02	1.40E-01	2pm	8.90E-02	1.32E-01	2pm	1.90E-01	3.30E-01	2pm	2.09E-01	3.05E-01	2pm	1.33E-01	3.67E-01	2pm	1.88E-01	3.62E-01	2pm
dsec16	9.06E-02	1.43E-01	2pm	9.33E-02	1.34E-01	2pm	1.97E-01	3.99E-01	2pm	2.38E-01	3.41E-01	2pm	1.54E-01	3.72E-01	2pm	2.02E-01	3.64E-01	2pm
dsec17	9.42E-02	1.37E-01	2pm	9.60E-02	1.34E-01	2pm	2.35E-01	4.38E-01	2pm	2.76E-01	3.79E-01	2pm	1.43E-01	3.90E-01	2pm	2.04E-01	3.67E-01	2pm
dsec18	1.29E-01	1.82E-01	2pm	1.32E-01	1.66E-01	2pm	2.90E-01	6.17E-01	2pm	3.97E-01	4.75E-01	2pm	1.90E-01	4.30E-01	2pm	2.47E-01	4.26E-01	2pm

dsec19	1.12E-01	1.62E-01	2pm	1.15E-01	1.60E-01	2pm	3.34E-01	3.78E-01	2pm	3.36E-01	3.77E-01	2pm	1.50E-01	5.14E-01	2pm	2.82E-01	4.55E-01	2pm
dsec20	1.86E-01	2.22E-01	2pm	1.87E-01	2.22E-01	2pm	2.94E-01	5.31E-01	2pm	3.50E-01	5.30E-01	2pm	2.51E-01	5.51E-01	2pm	3.41E-01	5.50E-01	2pm
dreg1	2.24E-01	2.85E-01	2pm	2.28E-01	2.08E-01	h	5.30E-01	6.48E-01	2pm	5.44E-01	4.62E-01	h	1.51E-01	2.57E+00	2pm	5.99E+00	1.32E+00	h
dreg2	2.00E-01	1.82E-01	h	2.00E-01	1.78E-01	h	4.64E-01	3.97E-01	h	4.69E-01	3.86E-01	h	1.08E-01	1.21E+00	2pm	1.32E+00	8.43E-01	h
dreg3	2.47E-01	2.26E-01	h	2.47E-01	1.99E-01	h	8.27E-01	7.28E-01	h	8.37E-01	5.63E-01	h	1.47E-01	1.46E+00	2pm	1.86E+00	9.41E-01	h
dreg45	2.75E-01	2.78E-01	2pm	2.75E-01	2.43E-01	h	4.47E-01	8.16E-01	2pm	5.83E-01	6.27E-01	2pm	2.98E-01	1.61E+00	2pm	2.01E+00	1.04E+00	h
dreg6	2.42E-01	2.89E-01	2pm	2.44E-01	2.36E-01	h	6.96E-01	5.71E-01	h	7.12E-01	5.65E-01	h	1.54E-01	2.86E+00	2pm	7.50E+00	1.42E+00	h
dreg7	2.22E-01	2.75E-01	2pm	2.25E-01	2.14E-01	h	5.43E-01	6.82E-01	2pm	5.63E-01	5.14E-01	h	1.17E-01	2.51E+00	2pm	5.86E+00	1.29E+00	h
dreg8	2.09E-01	1.70E-01	h	2.11E-01	1.66E-01	h	4.64E-01	3.10E-01	h	4.87E-01	3.10E-01	h	1.33E-01	1.52E+00	2pm	2.06E+00	9.36E-01	h
dreg9	1.84E-01	1.53E-01	h	1.84E-01	1.48E-01	h	3.88E-01	2.59E-01	h	4.04E-01	2.59E-01	h	8.40E-02	1.49E+00	2pm	2.05E+00	8.97E-01	h
dreg10	1.92E-01	2.15E-01	2pm	1.92E-01	1.74E-01	h	3.94E-01	5.05E-01	2pm	4.06E-01	3.78E-01	h	1.14E-01	1.76E+00	2pm	2.81E+00	1.01E+00	h
dreg11	3.12E-01	4.66E-01	2pm	3.36E-01	3.76E-01	2pm	8.18E-01	3.21E+00	2pm	6.52E+00	2.63E+00	h	2.65E-01	2.19E+00	2pm	3.96E+00	1.31E+00	h
dreg12	2.37E-01	2.06E-01	h	2.38E-01	1.79E-01	h	4.88E-01	3.37E-01	h	5.11E-01	3.27E-01	h	1.68E-01	2.49E+00	2pm	5.57E+00	1.25E+00	h
dreg13	1.88E-01	2.00E-01	2pm	1.89E-01	1.65E-01	h	4.00E-01	4.94E-01	2pm	4.09E-01	3.66E-01	h	1.01E-01	1.61E+00	2pm	2.39E+00	9.51E-01	h
dreg14	2.82E-01	2.95E-01	2pm	2.82E-01	2.33E-01	h	7.69E-01	1.17E+00	2pm	9.28E-01	8.05E-01	h	1.71E-01	2.08E+00	2pm	3.81E+00	1.15E+00	h
dreg15	1.99E-01	2.03E-01	2pm	1.99E-01	1.94E-01	h	5.21E-01	5.11E-01	h	5.21E-01	4.86E-01	h	1.02E-01	1.22E+00	2pm	1.35E+00	8.64E-01	h
dreg16	1.92E-01	1.81E-01	h	1.92E-01	1.65E-01	h	4.26E-01	3.64E-01	h	4.30E-01	3.29E-01	h	9.97E-02	1.70E+00	2pm	2.65E+00	9.84E-01	h
dyear	5.17E-03	5.83E-03	2pm	5.17E-03	5.41E-03	2pm	1.32E-02	1.48E-02	2pm	1.32E-02	1.44E-02	2pm	8.10E-03	2.93E-02	2pm	8.55E-03	1.72E-02	2pm
cons	3.66E-01	1.96E+00	2pm	2.92E+00	8.00E-01	h	7.66E-01	9.30E+00	2pm	7.36E+01	4.42E+00	h	5.26E-01	2.22E+00	2pm	3.40E+00	1.68E+00	h

A.3. Decomposition for the Two-Part model. Detailed and Complete Results

Table A.4. Decomposition for the two-part model. Estimation without firm-specific effects. Detailed and complete results

	2001				**2002**			
	participation eq		quantity eq		participation eq		quantity eq	
Training differential	**0.40143463**		**0.18120887**		**0.45014563**		**0.28994218**	
	Charact	Impact	Charact	Impact	Charact	Impact	Charact	Impact
Total	0.461	-0.059	0.25	-0.069	0.461	-0.011	0.248	0.041
	114.87%	-14.87%	138.14%	-38.14%	102.46%	-2.46%	85.55%	14.45%
White collars	0.018	-0.021	-0.004	-0.004	0.021	-0.006	0.013	-0.004
	4.57%	-5.30%	-2.40%	-1.97%	4.64%	-1.48%	4.55%	-1.31%
Advanced Technology	-0.04	-0.017	0.002	0.213	-0.041	-0.01	0.005	0.212
	-9.93%	-4.07%	0.87%	117.52%	-9.06%	-2.12%	1.76%	72.92%
Innovation	0.071	0.001	0.058	0.045	0.065	0.003	0.031	0.026
	17.68%	0.17%	31.90%	24.92%	14.53%	0.73%	10.81%	8.80%
International Market	0.051	-0.009	0.034	0.036	0.054	-0.009	0.035	0.034
	12.73%	-2.26%	18.98%	20.18%	11.95%	-1.97%	12.23%	11.63%
Foreign capital	0.025	-0.006	0.054	-0.109	0.027	-0.002	0.064	-0.101
	6.16%	-1.44%	29.77%	-60.35%	5.95%	-0.59%	22.24%	-34.69%
Temporary workers	0.001	-0.007	0.005	-0.202	0.001	-0.003	0.021	-0.177
	0.19%	-1.76%	2.73%	-111.38%	0.19%	-0.70%	7.12%	-61.12%
Size	0.258	-0.05	-0.055	0.947	0.265	0	-0.055	0.957
	64.39%	-12.64%	-30.24%	522.96%	58.95%	0.11%	-18.90%	329.81%
Productive capacity	-0.001	-0.009	0	-0.023	-0.002	-0.005	0	-0.024
	-0.34%	-2.32%	-0.01%	-12.51%	-0.39%	-1.07%	-0.03%	-8.36%
Group	0.031	-0.008	0.111	0.021	0.032	-0.005	0.111	0.023
	7.73%	-2.04%	61.11%	11.25%	7.16%	-1.17%	38.42%	8.07%
Sector	0.009	-0.037	0.029	0.188	0.002	-0.005	0.016	0.177
	2.31%	-9.15%	16.09%	103.52%	0.51%	-1.15%	5.58%	60.74%
Region	0.038	-0.348	0.017	-1.699	0.036	0.068	0.005	-1.646
	9.39%	-86.82%	9.34%	-937.56%	8.04%	14.95%	1.77%	-567.71%
Year	0	-0.019	0	-0.05	0	0	0	0
	0%	-4.64%	0%	-27.81%	0%	0%	0%	0%

Table A.5. Decomposition for the two-part model. Estimation including firm-specific effects. Detailed and complete results

	2001				**2002**			
	participation eq		quantity eq		participation eq		quantity eq	
Training differential	**0.49120915**		**0.22525114**		**0.54167439**		**0.27912344**	
	Charact	Impact	Charact	Impact	Charact	Impact	Charact	Impact
Total	0.571	-0.079	0.238	-0.012	0.578	-0.036	0.22	0.059
	116.26%	-16.26%	105.47%	-5.47%	106.71%	-6.71%	78.97%	21.03%
White collars	0.026	-0.03	-0.004	-0.029	0.03	-0.023	0.013	-0.032
	5.21%	-6.21%	-1.91%	-12.74%	5.47%	-4.16%	4.67%	-11.25%
Advanced Technology	-0.057	-0.035	-0.009	0.208	-0.06	-0.028	-0.005	0.206
	-11.68%	-7.16%	-4.13%	92.31%	-11.02%	-5.12%	-1.84%	73.82%
Innovation	0.067	0	0.042	0.061	0.063	0.003	0.023	0.035
	13.70%	0.08%	18.76%	27.06%	11.64%	0.45%	8.21%	12.30%
International Market	0.068	-0.014	0.042	0.036	0.073	-0.014	0.043	0.033
	13.87%	-2.80%	18.50%	15.82%	13.46%	-2.62%	15.39%	11.80%
Foreign capital	0.028	-0.008	0.052	-0.1	0.031	-0.006	0.063	-0.092
	5.73%	-1.59%	23.28%	-44.45%	5.72%	-1.05%	22.46%	-33.14%
Temporary workers	0	-0.005	0.005	-0.228	0	-0.004	0.023	-0.201
	0.05%	-0.99%	2.43%	-101.39%	0.05%	-0.61%	8.18%	-72.18%
Size	0.332	-0.112	-0.076	1.191	0.346	-0.064	-0.076	1.202
	67.52%	-22.92%	-33.66%	528.79%	63.88%	-11.89%	-27.17%	430.45%
Productive capacity	-0.001	-0.028	-0.001	-0.165	-0.002	-0.025	-0.006	-0.163
	-0.25%	-5.88%	-0.49%	-73.04%	-0.29%	-4.61%	-2.01%	-58.37%
Group	0.048	-0.012	0.103	0.027	0.051	-0.01	0.103	0.029
	9.83%	-2.44%	45.61%	11.82%	9.41%	-1.80%	37.02%	10.39%
Sector	0.011	-0.03	0.029	0.245	-0.002	-0.015	0.016	0.234
	2.34%	-6.21%	12.97%	108.79%	-0.41%	-2.86%	5.79%	83.94%
Region	0.049	-0.406	0.054	-1.661	0.048	-0.159	0.023	-1.607
	9.95%	-82.70%	24.11%	-737.15%	8.82%	-29.35%	8.26%	-576.06%
Year	0	-0.021	0	-0.014	0	0	0	0
	0%	-4.26%	0%	-6.33%	0%	0%	0%	0%