DECOMPOSING DIFFERENCES IN TOTAL FACTOR PRODUCTIVITY ACROSS FIRM SIZE

Laia Castany, Enrique López-Bazo, Rosina Moreno

AQR-IREA
Universitat de Barcelona, Avda Diagonal 690, 08034 Barcelona
Tel: +34 934021010   +34 934037041   +34 934027042   FAX: +34 934021821
Email: lcastany@ub.edu; elopez@ub.edu; rmoreno@ub.edu

Abstract: This paper investigates the extent to which the gap in total factor productivity between small and large firms is due to differences in the endowment of factors determining productivity and to the returns associated with these factors. We place particular emphasis on the contribution of differences in the propensity to innovate and in the use of skilled labor across firms of different size. Empirical evidence from a representative sample of Spanish manufacturing firms corroborates that both differences in endowments and returns to innovation and skilled labor significantly contribute to the productivity gap between small and large firms. In addition, it is observed that the contribution of innovation to this gap is caused only by differences in quantity, while differences in returns have no effect; in the case of human capital, however, most of the effect can be attributed to increasing differences in returns between small and large firms.

Keywords: Total Factor Productivity; skilled labor; innovation; firm size; Oaxaca decomposition

1. Introduction

The relationship between size and productivity has been studied for some time now both at a theoretical and empirical level. Jovanovich (1982) proposed a model in which, as firms gain experience, they learn about their levels of costs and efficiency. If they are efficient, they will decide to expand, or otherwise, they will decide to contract or even exit the market. Thus, the most efficient firms are expected to survive and grow, while the most inefficient will fail. In addition, Ericson and Pakes (1995) propose a model in which firms invest in R&D to improve their productivity levels. This way, their productivity is a function of their own R&D investment, of the productivity of their competitors and of the pressure of firms entering the market. If a firm succeeds and is sufficiently productive, it will grow. In the model by Olley and Pakes (1996), firms will decide to continue in the market and demand a certain amount of inputs, or otherwise, exit the market. This decision depends on whether firms expect to achieve a certain (unobservable) efficiency level or not. As long as firms continue in the market, firm size and productivity are also related in this model.

On the basis of these ideas, it is sensible to think that controlling for firm size in regressions to analyse firm productivity can even be considered a routine procedure (Geroski, 1998). Thus, when we add a firm size variable to a regression we are accounting for different technologies associated with a certain firm size, which, according to Geroski, is referred to as the direct effect of size on productivity: that is, as a variable that *ceteris paribus* improves efficiency. This author claims that size may also have an indirect effect on productivity, conditioning the effect of other variables on productivity as they will show different patterns of behaviour for small and large firms.

Technological and human capital are among the main factors that have traditionally been considered to foster productivity. Griliches (1979) is a pioneer study in assessing the contribution of research and development to productivity growth. Most literature on innovative activity estimates the elasticity or the rates of return to a stock of knowledge (calculated on the basis of the R&D effort) on productivity. Studies that use firm level data show a wide range of estimates and some have found weaker correlations than at sectoral or national level, particularly when industry dummies are included (see Mairesse and Sassenou, 1991, for a survey). However, the relationship between productivity and R&D expenditure embodies two different processes: the production of innovations starting from R&D activities, and the incorporation of these innovations into production (Griffith *et al.*, 2004). Firms invest in R&D in order to develop process and product innovations, which in turn may contribute to their productivity and economic performance. Crépon *et al.* (1998) emphasize that it is not

innovative input (R&D) but rather innovative output that increases firm productivity. Measures of innovative output allow us to measure the changes that firms consider relevant to their production processes and avoid the need to distinguish between formal and informal R&D activities (Huergo and Jaumandreu, 2004a).

Other studies have gone a step further by relating the innovative capacity to firm characteristics such as size and have found positive relation between them. Schumpeter (1942) hypothesized that large firms have an advantage over small companies as their financial situation allows them to be the most capable innovators. Acs *et al*. (1994) found that large firms invest more in R&D and innovate more, although small firms appear to have higher innovative productivity. Huergo and Jaumandreu (2004b) in Spanish case study reach the conclusion that "innovation is strikingly related to size". In these studies the underlying hypothesis is that firms show different patterns of innovative activity according to size. However, small and large firms can differ not only in their innovative intensity but also in the returns to such innovative activity. Klepper (1996) proposed a theoretical model in which firm size plays a crucial role in the appropiation of returns to innovation and in the engagement in R&D activities. The larger the firm, the greater the output over which fixed costs of process R&D can be averaged, which means that returns to process innovations are higher and encourages additional innovative effort. Cohen and Klepper (1996) corroborate the hypothesis of easier appropiability of returns to innovation in the case of large firms for the US case. These papers suggest the existence of differences in returns to innovation between small and large firms. Parisi *et al*. (2002) analyze the impact of innovations on productivity for different firm sizes for the Italian case and conclude that innovation in large firms has a greater impact on productivity.

On the other hand, Becker (1964) is a pioneering study in highlighting the importance of human capital. The literature on the effects of human capital on productivity argues that those workers with better problem solving and communication skills will perform any task requiring more than simple labor in a more efficient way. The microeconomic literature, which concerns the impact of investment in human capital on productivity levels, has typically estimated Mincerian equations (see Harmon *et al*. 2003, for a survey). Nevertheless, only a few microeconomic studies have analyzed the effect of human capital on productivity at the firm level. One exception is Griliches and Regev (1995), which estimates a production function that includes R&D capital services and a measure of labor quality as a proxy for human capital in Israeli industry. The authors find a coefficient of skilled labor of approximately 0.4 for the total sample and 0.5 for large firms in the pooled regressions using

differences to control for individual heterogeneity, which they acknowledge to be quite high. Haltiwanger *et al*. (1999) used a matched employer-employee dataset and found that labor productivity is associated with certain characteristics of the workforce, such as the proportion of educated workers. The results are consistent with a human capital model in which more highly-skilled workers make the firm more productive. In addition to the fact that large firms employ a more skilled labor force, it is also possible that returns to human capital are higher in large firms. In fact, in the literature that aims to explain the positive relationship between firm size and wages, Oosterbeek and Van Praag (1995) found that larger firms pay higher returns to human capital (which can explain part of the size-wage relation).

The empirical evidence and the arguments given above suggest that technological and human capital (which can be combined under the term *knowledge capital*) affect productivity. There are also arguments which suggest that larger firms have better endowments of these factors. We argue that technological and human capital play different roles in determining productivity for small and large firms: firstly, because large firms are usually more innovative and employ more skilled workers; secondly, because the returns of these endowments on productivity may be higher in the case of large firms. In other words, an additional innovation implemented in a large firm or a skilled employee hired by the firm would result in a higher return than if the same situation arose in a small firm. Consequently, size would be exerting an indirect influence on firm productivity, as it conditions the impact of other factors on productivity. According to this hypothesis, the observed gap in productivity between small and large firms would be caused by both differences in the endowment of technological and human capital between firms of different size, and by differences in the returns to the use of these types of capital.

Therefore, the main objective of this paper is to investigate the differences in total factor productivity (TFP) between small and large firms and to establish the extent to which this gap is caused by differences in the endowment of the main factors that determine firm productivity or to the returns obtained from these factors. Building on this idea, one of the main contributions of the present analysis is the use of the traditional Oaxaca-Blinder decomposition to assess the relative importance of differences in firm endowments and in their returns to explain the productivity differential between small and large firms. To the best of our knowledge, Smith *et al*. (2004) is the only study that uses this decomposition to analyze differences between firms. More specifically, the authors compare firms that carry out R&D activities and firms that do not.

We provide empirical evidence for a representative sample of Spanish manufacturing firms in the period from 1990 to 2002. The results corroborate that both differences in endowments and returns to knowledge capital significantly contribute to the TFP gap between small and large firms. In addition, it is observed that the contribution of innovation to this gap is caused only by differences in quantity and that differences in returns have no effect, while in the case of human capital the contribution is largely attributable to increasing differences in returns between small and large firms.

The paper is structured as follows. After the introduction, Section 2 presents the empirical specification and some methodological issues. In Section 3 we present the data and make a descriptive analysis. Section 4 shows the results and Section 5 concludes the study.

2. Empirical specification and methodological issues

Our empirical framework relates a TFP index to innovation and skilled labor, the variables of interest and to several control variables. Our approach is fairly close to that of Griliches and Regev (1995), who estimate production function at firm level including measures of human and technological capital. Instead of the production function, we use an estimate of the TFP as the dependent variable and innovation and skilled labor as the explanatory variables whose effects on productivity we wish to assess.[1] The empirical model can therefore be expressed as follows:

$$TFP_{it} = \beta_0 + \beta_1 INN_{it} + \beta_2 H_{it} + Z'\gamma_{it} + u_{it} \qquad (1)$$

where TFP is the logarithm of the total factor productivity index in firm i in year t, INN and H are variables proxying for the innovation activity and the level of human capital in the firm, Z is a set of standard control variables (size, age, industry and year effects) and u is an error term.

In order to analyze the causes of differences in TFP by firm size we use the methodology given by the Oaxaca-Blinder decomposition. This decomposition has been

---

[1] We overcome the endogeneity problem in labor, capital and materials when estimating production functions by calculating a TFP index and using input prices instead of estimating their returns to calculate the participation of each input in the production function. Endogeneity problems associated with the demand of labor, capital and intermediate inputs when estimating production functions are well known: the demands of inputs not only determine productivity, but they also depend on the productivity they obtain. Then, the residual is correlated with the part of the inputs that is endogenously determined, which produces biased coefficients for the inputs and therefore an inconsistent estimation of the production function parameters. In addition, if a given variable is omitted in the estimation of the production function and this variable is also relevant, the error term in the production function and the demand of inputs will be correlated, which produces biased coefficients.

widely used to study wage gaps associated with differences in worker characteristics and with discrimination by gender or race.[2] We apply the decomposition to analyze differences in TFP between small and large firms. This methodology allows us to analyze the extent to which the TFP differential between small and large firms can be explained by differences in *endowments* of human and technological capital and differences in the *returns* to these endowments.

The standard methodology decomposes the TFP differential between small and large firms based on two auxiliary regressions specified as in (1) for each type of firm. From these regressions, the average TFP in the sample of small and large firms is obtained as:

$$\overline{TFP}_S = \overline{X}_S{}'\hat{\beta}_S$$
$$\overline{TFP}_L = \overline{X}_L{}'\hat{\beta}_L \qquad (2)$$

where $\overline{TFP}$ denotes the mean of TFP, $\overline{X}$ is the vector of the mean values of the regressors in specification (1), $\hat{\beta}$ is the conforming vector of estimated coefficients and subscripts S and L refer to small and large firms respectively. Then, differences in TFP between small and large firms can be decomposed as:

$$\overline{TFP}_L - \overline{TFP}_S = \left(\overline{X}_L{}' - \overline{X}_S{}'\right)\hat{\beta}_L + \overline{X}_S{}'\left(\hat{\beta}_L - \hat{\beta}_S\right) \qquad (3)$$

where the first term on the right-hand side is the part of the TFP gap caused by differences in characteristics between the representative small and large firms, and the second term on the right-hand side is the contribution of differences in returns between the two types of firm.

The first term on the right hand side of expression (3) assumes that all the firms have the returns of large firms, $\hat{\beta}_L$. In the second term, we assume that all firms have the endowments of small firms, $\overline{X}_S$. However, we could write a symmetric equation in which these values are replaced by $\hat{\beta}_S$ and $\overline{X}_L$ respectively. The standard version of the decomposition is based on the assumption that one of the two equations is the "natural" one. For example, in the case of wage differentials by gender, it may appear fairly natural to assume that women are the "discriminated" group: we would assume that women have the same returns as men for the first term, and the second term could therefore be interpreted as wage discrimination by gender. In our case there is no compelling reason to calculate the differences in firm endowments with the assumption that all firms had the returns of either

---

[2] Following Oaxaca (1973) and Blinder (1973), the difference between the mean wages between two groups, for example men and women, in period *t* can be decomposed into an explained or predicted difference due to disparities in observed or measured characteristics between the two groups, and an unexplained or residual difference attributable to both wage discrimination and unmeasured disparities in characteristics.

large or small firms. One strand of the literature considers that it is not always easy to establish which is the natural model and that results may often differ considerably. This literature suggests a variation of the standard decomposition in which it is not necessary to make an assumption about which is the natural model. According to this perspective, there is a "non-discriminatory structure of returns" in relation to which one group is "discriminated" while the other is "favored". The TFP differential without the assumption that one of the two equations is the natural model is expressed as:

$$\overline{TFP}_L - \overline{TFP}_S = \left(\overline{X}_L' - \overline{X}_S' \right) \hat{\beta}^* + \overline{X}_L' \left(\hat{\beta}_L - \hat{\beta}^* \right) + \overline{X}_S' \left(\hat{\beta}^* - \hat{\beta}_S \right) \tag{4}$$

where $\hat{\beta}^*$ is the estimated non-discriminatory returns structure. The first term on the right-hand side of (4) is an estimate of the productivity differential in the absence of differences in returns between small and large firms, which reflects productivity differences caused by differences in firm endowments. The second and third terms are estimates of the advantage of large firms and the disadvantage of small firms in relation to the non-discriminatory returns structure. The two terms together are considered differences in TFP according to firm size, associated with differences in returns without imposing a discriminated group. Since we are not interested in distinguishing the advantage and disadvantage effects but in evaluating the differences in returns as a whole, without imposing a discriminated returns structure, here we will report these two terms together.[3] To implement this decomposition, it is necessary to make an assumption regarding what the non-discriminatory returns structure would be. Oaxaca and Ransom (1994) deal with a proper selection of the non-discriminatory structure and propose estimating it as a weighted average of the two returns structures, $\hat{\beta}^* = \Omega \hat{\beta}_L + (I - \Omega) \hat{\beta}_S$, in which the weights Ω are calculated as $(X'X)^{-1}(X'_L X'_L)$ and where X is the matrix of regressors for the entire sample of firms.

3. Data and Descriptive Analysis

3.1. Description of the dataset

We use a sample of Spanish manufacturing firms from the official survey Encuesta de Estrategias Empresariales (ESEE). This survey is an unbalanced panel that covers the period 1990-2002 and collects information on strategic decisions and the behavior of firms. Firms answered a comprehensive questionnaire every four years (covering those issues that would supposedly change annually) and a reduced questionnaire in the intervening years, so that

---

[3] Detailed results are available from the authors upon request.

complete information is available for 1990, 1994, 1998 and 2002. The reference population of the ESEE was a sample of firms with 10 or more employees, working in one of the activities corresponding to divisions 15 to 37 of NACE-93, excluding division 23 (activities related to oil refinement and fuel treatment). During the initial period, all firms with more than 200 employees were required to participate (70% did). Firms with 10 to 200 employees were sampled randomly according to industry and four size strata, retaining about 5%, in order to guarantee representativity for every industry and firm size. The ESEE is designed to change as the composition of industry evolves. Newly created firms are selected using the original selection criteria. Due to death and attrition, some firms were replaced by others in their industry and size group so as to maintain representativity.[4]

The measure of TFP we use is based on the index developed by Good *et al*. (1996), which is derived from a translog production function. Its analytical expression for a firm *f* in period *t* is as follows:

$$\ln TFP_{ft} = (\ln Y_{ft} - \overline{\ln Y_t}) - \frac{1}{2}\sum_{i=1}^{n}(S_{ift} + \overline{S_{it}})(\ln X_{ift} - \overline{\ln X_{it}})$$

$$+ \sum_{s=2}^{t}(\overline{\ln Y_s} - \overline{\ln Y_{s-1}}) - \frac{1}{2}\sum_{s=2}^{t}\sum_{i=1}^{n}(\overline{S_{is}} + \overline{S_{i,s-1}})(\overline{\ln X_{is}} - \overline{\ln X_{i,s-1}})$$

where *Y* and X_i are quantities of output and the *i*-th input respectively, S_i is the cost-based share of the *i*-th input and the bar over variables refers to their average. The upper part of the expression is the deviation of the firm output and inputs from those of a hypothetical firm, which is the reference point in year *t*. The lower part of the expression is the cumulative change in the output and input reference point between year *t* and the initial year. This second part introduces a productivity differential every year (as output, inputs and shares may change) and therefore accounts for possible technological changes. The productivity index for a given firm and year is expressed in relation to the hypothetical firm in the base period, 1990. Following Hall's (1990) suggestion, weights are calculated as the share of every input in the total cost of inputs.

Firm size *(SIZE)* is defined as the log of the total number of employees. In our database, small firms are those with 200 or fewer employees. Many databases and studies consider that small and medium enterprises are those firms with fewer than 250 employees. However, our database makes the distinction at 200 employees and uses different sampling

---

[4] See Fariñas and Jaumandreu (1999) for further details.

schemes for the two groups. We consider it is more appropriate to use the same criterion to guarantee representativity by size strata.[5]

The literature suggests a wide variety of variables for measuring innovative activity at the firm level. On the one hand, R&D expenditure is a measure of innovative inputs or the R&D effort of firms. On the other hand, the innovative capacity can also be measured by process and product innovations. These variables are a measure of the innovative output or the innovative effort that effectively becomes an innovation. Similarly to previous studies of Spanish manufacturing firms (Huergo and Jaumandreu 2004a), we use process innovation as the measure of a firm's innovative activity, as it is assumed that process and not product innovations are the ones that improve the mechanisms through which inputs are transformed into output (Ornaghi, 2006). Specifically, it is a dichotomous variable (*INN*) that takes value 1 if the firm has implemented a process innovation in the same year. Process innovations improve the mechanisms through which an input is transformed into output. Process innovations are assumed to take place when the firm gives a positive response to the following request: "Indicate if your firm introduced some significant modification in the production process (process innovation). If the answer is yes, please indicate the way: (a) introduction of new machines; (b) introduction of new methods of organization; (c) both".

Human capital (*HK*) is measured in terms of formal education of the labor force. This variable is defined as the ratio of skilled workers according to educational level. Specifically, it includes engineers, graduates, middle level engineers, experts and qualified assistants.

Of the control variables, the age (*AGE)* is defined as the number of years since the firm was set up, whereas the sectoral heterogeneity is proxied by a set of 20 dummy variables (*DSEC)* according to the NACE-93 classification, where the omitted category is "Other manufacturing industries". Finally, a set of year dummies is included to control for exogenous technical progress and effects of the business cycle that are common to all firms (*DYEAR*).

3.2. Descriptive Analysis

Table 1 shows some basic statistics for TFP for the years on which our analysis is focussed (1994, 1998 and 2002) for the total sample, and for the subsamples of small and large firms.[6] The evolution of TFP over the period under analysis in our sample is similar to that reported

---

[5] Delgado *et al.*(2002), Fariñas and Ruano (2004), and Máñez *et al*. (2004) use the same criterion when using data from the ESEE.

[6] As explained later in the paper, we use the lag of innovative activity and the percentage of skilled workers to mitigate the effects of simultaneity. Therefore, we give the description of these measures for only 1994, 1998 and 2002, since the description for 1990 cannot be considered in the estimation.

at the firm and aggregate levels in some other studies of the Spanish economy (Estrada and López-Salido, 2001; Huergo and Moreno, 2006). If we focus on TFP differences by firm size, we find that large firms have higher TFP than small ones and that the difference is statistically significant for all three years. In fact, the t-test for equality of means rejects the null hypothesis that the average TFP is equal in the two subsamples.

[Insert Table 1 around here]

Tables 2 and 3 contain basic descriptive statistics for the two other variables of interest: innovation and human capital. The percentage of skilled workers increases over time but, surprisingly, the firms do not report an increase in innovative output between 1994 and 2002. If we focus on differences by firm size, these tables confirm our *a priori* reasoning that large firms are associated with greater innovation activity and a more skilled labor force. The figures show that the proportion of innovative large firms almost doubles that of innovative small firms. In terms of human capital, large firms also report much larger proportions of skilled workers than small firms. The test for equality of means rejects the null that, on average, the two types of firm have the same ratio of white-collar workers, and the test for equality of proportions rejects the null that small and large firms report similar proportions of innovation activity. Differences in innovation and skilled labour endowments between small and large firms remain quite stable over time, which is reflected later in the decomposition of the TFP gap.

[Insert Tables 2 and 3 around here]

This descriptive analysis so far basically shows that large firms are more productive, more innovative and invest more in human capital than their smaller counterparts. In addition, Table 4 shows that innovation and employing skilled workers are associated with higher TFP levels in both small and large firms. The table shows the average TFP for innovative and non-innovative firms with different proportions of skilled workers and considering small and large firms separately. In general, TFP increases as the proportion of white-collar workers increases, which indicates a positive relationship between these variables. In general, small innovative firms in a given quartile of human capital distribution have higher TFP than small non-innovative firms. The same result is found for large innovative and non-innovative firms. When we compare small innovative and large innovative firms, we also find that the large firms are more productive than their smaller counterparts. Finally, the large non-innovative

firms are more productive than small non-innovative firms with the same percentage of skilled workers. These results form the basis of the following analysis. Firms that innovate and have a higher proportion of skilled workers are more productive, and among these firms, large firms are also more productive.

[Insert Table 4 around here]

4. Results

4.1. Estimation of the TFP specifications

The first step in our analysis is to estimate by OLS the empirical specification in (1) for the total sample of firms and for the small and large sub-samples. It should be noted that, to mitigate the effects of simultaneity, we use the lag of innovative activity and the percentage of skilled workers. For the total sample, column (i) of Table 5 shows that the estimated coefficients for the two variables of interest (innovative activity and worker qualifications) are positive and significant at 1%. Therefore, our results support the hypothesis that the knowledge capital used by a firm improves the mechanism through which inputs are transformed into output. Specifically, process innovations reduce the unitary cost of production and contribute to increased productivity. However, the effect seems to be modest: changing from a non-innovative to an innovative firm increases TFP by only 3.4%.

The contribution of human capital to productivity is also positive and clearly significant. The estimated coefficient for this factor confirms that more intensive use of skilled labor produces a higher level of productivity. Specifically, increasing the ratio of skilled workers by 10% raises the TFP of the average Spanish manufacturing firm by 2.1%. For the control variables, both firm size and age have a significant coefficient with the expected sign, while the sets of industry and time dummies are jointly significant.

[Insert Table 5 around here]

Since we are interested in studying whether innovation and human capital make different contributions to productivity enhancement in firms of different size, we estimate their effect on the sub-samples of small and large firms separately. As shown in columns (ii) and (iii) of Table 5, coefficients for innovation are positive and significant at 1% in both cases. More concretely, small and large innovative firms are more productive than their respective non-innovative counterparts. However, the difference in the effect of innovation on

firms of different size is moderate (3.5% for small firms versus 3.7% for large firms), which suggests that once the innovation is implemented it exerts the same effect on productivity regardless of firm size.

The effect of human capital is also positive and significant at 1% in both types of firm. However, the most interesting feature of this variable is that its effect on TFP seems to be much greater for large firms. Specifically, increasing the ratio of skilled workers by 10% raises TFP by 3% in large firms and by only 1.8% in the case of small firms. Coefficients of the control variables remain significant for both small and large firms, except for the coefficient of size in the sample of large firms.

The general conclusion we can draw is that innovation and employing more skilled workers enhance productivity in both small and large firms. Furthermore, the impact of these variables on TFP seems to be greater in large firms than in small firms. The fact that the impact of human capital is different between the two sub-samples indicates that small and large firms do not only have different human capital endowments but also different returns to this factor. Therefore, the incentive to hire skilled workers as a way of increasing productivity seems to be stronger in large firms. In contrast, the differences in returns to innovative activity are not as large. This suggests that the impact of innovation on differences in TFP across firms of different size seems essentially to be related to differences in the propensity to innovate, which is much higher in large firms.

4.2. Robustness Analysis

Additional control variables

The results presented so far are based on our baseline specification in Equation (1), which controls for firm size and age, sector and time period. An initial robustness analysis of the estimated effects of knowledge capital (i.e. innovation and human capital) and differences between small and large firms is performed by including additional control variables related to particular characteristics of the firm. Specifically, the variables added to the initial specification are defined as follows:

- The proportion of the productive capacity used by the firm (*PRODCAP*).

- The proportion of foreign-owned capital (*FOREIGNK*), the proportion of publicly-owned capital (*PUBLICK*) and a dummy to represent whether the firm belongs to a group of firms (*GROUP*). They are intended to proxy for the ownership structure of the firm.

- The degree of competition faced by the firm, proxied by the geographical scope of its principal market (*MARKET*). We use a set of dummy variables to consider whether the

market is local and provincial, regional, national and international, or all three. The latter is considered as the omitted category.

- The (log of the) value of exports expressed in constant pesetas, base year 1990 (*EXPORT*).

- A set of 17 dummy variables for the Spanish regions as defined in the NUTS II database (*DREG*). The omitted category is "La Rioja".

It should be noted that the effect of knowledge capital may well be related to these additional control variables. Some characteristics may cause firms to obtain a higher return to the use of knowledge capital, so that when controls are included that proxy for these characteristics we obtain a lower estimate of the return.[7]

The results are summarized in columns (iv) to (vi) of Table 5. In general terms, the effect of the control variables is as expected, although whether the firm is supported by public capital and whether it belongs to a group do not seem to exert a significant effect on TFP over the whole sample. The effect of foreign capital and the volume of exports varies between small and large firms. However, more importantly for our analysis, the estimates of the effects of innovation and human capital are fairly robust to the inclusion of these additional control variables. The greatest change is observed in the estimates of the return to human capital which, as expected, decrease in magnitude (from over 20% to 14% over the whole sample, and a decrease of the same order of magnitude for the sample of small and large firms).

Unobserved heterogeneity

As well as introducing additional control variables, we estimate a random effects model to account for the existence of unobserved firm heterogeneity that is likely to affect our estimates of the returns to knowledge capital.[8] Columns (vii) to (ix) of Table 5 show the results of the estimation of the random effects model for the whole sample and for small and large firms. In all three samples, a standard Lagrange multiplier test clearly rejects the null of the absence of unobserved random effects.

When performing estimates with the whole sample, the coefficients of our variables of interest remain positive and significant at 5% after controlling for firm-specific heterogeneity

---

[7] This argument is similar to the one usually formulated when analyzing the return to human capital from the estimation of Mincerian wage equations (Pereira and Martins, 2004).

[8] This model assumes that firm heterogeneity is part of a compound error term that is uncorrelated with the regressors. It has been common practice to estimate a random effects model when, as in our dataset, firms in the sample are selected randomly from a larger population. Barrios *et al.* (2003), Máñez *et al.* (2004), and Licandro *et al.* (2004) use this type of specification in analyses using ESEE.

and the set of additional control variables. However, in this case the estimate of the two coefficients is lower than the estimate obtained when not controlling for unobserved heterogeneity. However, when firms are separated by size, the estimated effect of innovation and human capital only seems to be significant (at the usual probability levels) for large firms. In fact, the main conclusions are maintained in the case of large firms, with an estimated return to innovation of approximately 3.5% and a return to human capital of approximately 16.2%. However, neither of these coefficients is significant in the case of small firms.

In summary, both innovation and human capital seem to have an effect on productivity enhancement, although the evidence suggests that the magnitude of this effect is very strongly related to firm size. After controlling for an extensive set of conditioning variables and for unobserved firm heterogeneity, the effect of the knowledge capital variables is only marginal and is not statistically significant for the group of small firms.[9] This suggests the existence of a threshold in the effect of innovation and human capital that is related to the size of the firm.

Consequently, small and large firms show different patterns of behavior in relation to these variables: they not only have different endowments, but the returns to these endowments differ substantially between firms. In this framework we argue that differences in TFP between small and large firms are associated with differences in endowments and with differences in the returns to these endowments. In the next section we analyze the relative contribution of these effects to productivity differences according to firm size.

4.3 Decomposition of the TFP gap between small and large firms

As described in Section 2, the detailed decomposition of the productivity differential between small and large firms allows us to assess the relative contribution of firm characteristics and the returns to these characteristics, as well as decompose the individual effect of each variable. The individual decomposition is particularly useful in this study, since we are interested in determining the contribution of both human and technological capital. We calculate the decomposition based on the specification with additional control variables with both the OLS and the RE estimates. Although in our opinion the RE estimation would be preferable given that it considers unobserved firm heterogeneity, we also obtain the

---

[9] It should be kept in mind that by controlling for the observable and unobservable heterogeneity in the last set of estimates, we are considering the channels through which knowledge capital may be exerting its influence on TFP. In such a case, the estimate of the returns from the specification that include firm heterogeneity should be considered as a lower bound. This can be particularly important in the group of small firms if only those firms with favorable characteristics are able to obtain profit from innovations and human capital.

decomposition based on the OLS estimation since it provides an exact decomposition whereas the one based on RE does not.[10, 11]

Table 6 shows the results for each of the years under analysis, based on the OLS estimation (Panel A) and the RE estimation (Panel B), for the specification that includes the whole set of control variables. The TFP differential between small and large firms was approximately 0.08 in 1994 and decreased to 0.06 in 2002. The decomposition based on OLS estimates shows that differences in endowments almost completely account for the TFP gap. Specifically, in 1994, differences in endowments accounted for approximately 95% of the differential, while differences in returns accounted for the remaining 5%. Between 1994 and 2002, the contribution of endowments increased and reached 104% in 2002. A contribution of differences in endowments with a value higher than 100% implies that differences in aggregate returns favored small firms in the last two years of the analysis (negative contribution of differences in returns).

However, this general result is an aggregation of the effects of all variables in the specification, thus positive and negative effects may be compensated. As stated above, the detailed decomposition allows us to assess the individual contributions of innovation and human capital. The results suggest that these two variables together, including differences in endowments and differences in returns, account for a large proportion of the TFP gap. Moreover, their contribution is seen to increase over the period under analysis: the variables account for 32% of the gap in 1994 and 44% in 2002. The joint contribution of the two variables under analysis is the combined effect in similar proportions of differences in endowments and in returns. However, the individual analyses for innovation and human capital reveal significant differences in the sources of their impact on the TFP gap. Innovation can account for approximately 10% of the gap due exclusively to the fact that large firms innovate more than small firms (difference in characteristics). However, most of the contribution of human capital is caused by the higher returns in large firms. In addition, the portion of the TFP gap that can be explained by human capital increases over the period studied (from approximately 21% in 1994 to as much as 32% in 2002) and is also largely

---

[10] In the RE model the transformed residuals have zero mean, but the residuals from the original specification do not. This prevents us from obtaining an exact decomposition of the TFP gap based on the RE estimates of the coefficients.

[11] The decomposition described below is robust to the choice of omitted category in the dummy variables in the TFP specification, as the estimates used to compute it have been obtained by imposing the parametric identification constraints suggested in Gardeazabal and Ugidos (2004). This is particularly important in the case of innovation, which is one of the variables that interests us in this study.

caused by an increase in the contribution of differences in returns between small and large firms.

[Insert Table 6 around here]

A similar situation is observed when the decomposition is obtained by using the estimate of the coefficients from the RE specification (Panel B in Table 6). Although – in agreement with the lower magnitude for the estimate of the coefficients – the contribution of the knowledge variables is slightly lower in this case, their impact on the TFP gap seems to be considerable and it increases over time (21.5% in 1994, 23.2% in 1998, and 28.7% in 2002). The major difference is related to the lower contribution of innovation, which is caused by the lesser impact of differences in its endowment (as the RE estimation assign a lower weight – estimated coefficient – to differences in innovation between small and large firms). The results derived from the OLS estimates for human capital seem to be robust to the inclusion of unobserved firm heterogeneity: this variable alone accounts for a large portion of the TFP gap, its contribution increases over time, and it is mostly caused by the different returns observed in small and large firms.

5. Conclusions

Among the factors behind the differences in productivity between small and large firms, previous literature has highlighted the greater endowment of certain elements in large firms. In this paper we suggest that the higher productivity levels achieved by large firms may be due not only to the greater endowment of these elements but also to the higher returns they obtain from their use. We focus on knowledge capital and specifically on innovation and the use of skilled labor. One of our main points is that differences in productivity between small and large firms could be caused not only by the greater endowment of these elements in large firms but also by the greater impact of innovations and the use of skilled workers in large firms.

To analyze this issue we propose to decompose the gap in total factor productivity between small and large firms into two components: one caused by differences in endowments and another caused by differences in returns. We apply the traditional Oaxaca-Blinder methodology to a representative sample of Spanish manufacturing firms and find that innovation and human capital together account for a significant part of the productivity differential between small and large firms. Interestingly, the sources of the separate

contribution of these two factors are quite different. The contribution of innovation is due mainly to differences between firms in the propensity to innovate. This is because large firms carry out more innovation activity but, once the innovation is implemented, both small and large firms obtain the same benefit from it. In contrast, the part of the differential caused by human capital is due not only to differences in the use of skilled labor (again, higher in large firms) but also to the higher return that large firms obtain from skilled employees, an effect that tends to increase over time.

According to our results, policies focused on stimulating innovation in small firms may enhance productivity, which in turn would reduce the productivity gap between small and large firms and – given the proportion of small firms in some economies – enhance aggregate productivity. However, it would be necessary to provide incentives not only to make a more intensive use of skilled labor, but also to ensure that small firms obtain the maximum returns from the use of this human capital. Accordingly, future research should attempt to shed more light on the reasons behind this different return obtained from skilled labor in small and large firms, and on the connection with the size-wage effect reported in labor economics literature (Oi and Idson, 1999).

References

Acs, Z.J., Audretsch, D.B. and Feldman, M.P., (1994). "R&D Spillovers and Recipient Firm Size", The Review of Economics and Statistics, vol. 76(2), 336-40.

Barrios, S. Görg, H. and Strobl, E. (2003), "Explaining Firms' Export Behaviour: R&D, Spillovers and the Destination Market", Oxford Bulletin of Economics and Statistics, Vol. 65, 475, September.

Becker, G. (1964), "Human capital. A theoretical and empirical analysis", New York. National Bureau of Economic Research. Columbia University Press.

Blinder, A.S. (1973), "Wage Discrimination: Reduced Form and Structural Estimates", The Journal of Human Resources 8(4), 436-455.

Cohen, W.M. and Klepper, S. (1996), "Firm Size and the Nature of Innovation within Industries: The Case of Process and Product R&D", Review of Economics and Statistics 78, 232-243.

Crépon, B., Duguet, E. and Mairesse, J. (1998), "Research, innovation, and productivity an econometric analysis at the firm level", NBER Working Paper, No. 6696, August.

Delgado, M.A, Fariñas, J.C. and Ruano, S. (2002), "Firm productivity and export markets: a non-parametric approach", *Journal of International Economics*, vol. 57, p. 397–422.

Ericson, R. and Pakes, A. (1995), "Markov perfect industry dynamics: a framework for empirical analysis", Review of Economic Studies, Vol. 62, No. 1, 53-82.

Estrada, A. and López-Salido, J. D. (2001), "Sectoral and aggregate technology growth", Banco de España - Servicio de Estudios, Documento de Trabajo n.º 0116.

Fariñas, J.C. and Jaumandreu, J. (1999), "Diez años de Encuesta sobre Estrategias Empresariales (ESEE)", Economía Industrial, n° 329, 29-42.

Fariñas, J.C. and Ruano, S. (2004), "The Dynamics of Productivity: A Decomposition Approach Using Distribution Functions," *Small Business Economics*, vol. 22, p. 237-251.

Gardeazábal, J. and Ugidos, A. (2004), "More on Identification in Detailed Wage Decompositions", The Review of Economics and Statistics, November 2004, 86(4), 1034-1036.

Geroski, P.A. (1998), "An Applied Econometrician's View of Large Company Performance", Review of Industrial Organization, Vol. 13, 271-293.

Good, D.H., Nadiri, M.I. and Sickles, R.C. (1996), "Index Number and Factor Demand Approaches to the Estimation of Productivity", NBER Working Paper Series, 5790.

Griffith, R., Redding, S. and Van Reenen, J. (2004), "Mapping the two faces of R&D: productivity growth in a panel of OECD industries", The Review of Economics and Statistics, 86(4), 883.895.

Griliches, Z. (1979), "Issues in Assessing the Contribution of Research and Development to Productivity Growth", Bell Journal of Economics, The RAND Corporation, vol. 10(1), 92-116.

Griliches, Z. and Regev, H. (1995), "Firm productivity in Israeli industry, 1979-1988", Journal of Econometrics, Vol. 65, 175-203.

Hall, R.E. (1990), "Invariance Properties of Solow's Productivity Residual", In Diamond (ed.), "Growth/Productivity/Employment", Cambridge, MA: MIT Press, 71-112.

Haltiwanger, J.C., Lane, J.I. and Spletzer, J.R. (1999), "Productivity Differences across Employers: The Roles of Employer Size, Age, and Human Capital", American Economic Review, vol. 89(2), 94-98, May.

Harmon, C., Oosterbeek, H. and Walker, I. (2003), "The Returns to Education: Microeconomics", Journal of Economic Surveys, Blackwell Publishing, vol. 17(2), 115-156.

Huergo, E. and Jaumandreu, J. (2004a), "Firms' age, process innovation and productivity growth", International Journal of Industrial Organization, Vol. 22, 541-559.

Huergo, E. and Jaumandreu, J. (2004b), "How does probability of innovation change with firm age?" Small Business Economics, 22, 193-207.

Huergo, E. and Moreno, L. (2006), "La productividad de las empresas manufactureras españolas en la década de 1990", in Segura, J. (Coord.) "La productividad en la economía española", Fundación Ramón Areces, Madrid.

Jovanovich, B. (1982), "Selection and the evolution of industry", Econometrica, Vol. 50, No. 3, May, 649-70.

Klepper, S., 1996, "Entry, Exit and Innovation Over the Product Life Cycle", American Economic Review, 86, pp 562-583.

Licandro, O., Maroto, R. and Puch, L. (2004), "Innovation, Investment and Productivity: Evidence 10:40 23/11/2006 From Spanish Firms." EUI Working Paper Eco No. 2004/7.

Mairesse, J. and Sassenou, M. (1991), "R&D and Productivity: A Survey of Econometric Studies at the Firm Level", NBER Working Paper, No. 3666.

Máñez, J. A., Rochina, M. E., Sanchis, A. and Sanchis, J. A. (2004), "A Dynamic Approach to the Decision to Invest in R&D: the role of sunk costs", mimeo.

Oaxaca, R. (1973), "Male-Female Wage Differentials" in Urban Labor Markets International Economic Review 14, 693-709.

Oaxaca, R., and Ransom, M. R. (1994), "On discrimination and the decomposition of wage differentials", Journal of Econometrics 61, 5.21.

Oi, W. and Idson, T. (1999), "Firm size and wages", in Ashenfelter, O. and Card, D. (eds) "Handbook of Labor Economics", vol IIIB, North-Holland, Amsterdam.

Oosterbeek, H. and van Praag, M. (1995), "Firm-Size Wage Differentials in the Netherlands", Small Business Economics, vol. 7(3), 173-82, June.

Olley, G.T. and A. Pakes (1996), "The dynamics of productivity in the telecommunications equipment industry", Econometrica, Vol. 64(6), 1263-1297.

Ornaghi, C. (2006), "Spillovers in product and process innovation: Evidence from manufacturing firms", International Journal of Industrial Organization 24 349- 380.

Parisi, M.L., Schiantarelli, F. and Sembenelli, A. (2002). "Productivity, Innovation Creation and Absorption, and R&D: Micro Evidence for Italy," Boston College Working Papers in Economics 526, Boston College Department of Economics.

Pereira, P.T. and Martins P.S. (2004) "Returns to education and wage equations", Applied Economics 36 (6) 525 -531.

Schumpeter, J.A. (1942), "Capitalism, Socialism and Democracy", New York, Harper and Brothers.

Smith, V., Dilling-Hansen, M., Eriksson, T. and Strojer-Madsen, E. (2004), "R&D and productivity in Danish firms: some empirical evidence", Applied Economics, 36, 1797-1806.

Temple, J. (2001). Generalizations that aren't? Evidence on education and growth. European Economic Review, May, 45(4-6), 905-918.

Table 1. Description of the TFP index in Spanish manufacturing firms

	1994			1998			2002		
	Total sample	Small firms	Large firms	Total sample	Small firms	Large firms	Total sample	Small firms	Large firms
Mean	0.0204	-0.0011	0.0782	0.0749	0.0586	0.1352	0.1184	0.1054	0.1657
Variance	0.0539	0.0578	0.0389	0.0661	0.0685	0.0529	0.0716	0.0763	0.052
Test Eq means		4.9087***			4.1136***			3.0421***	
No Obs	852	621	231	968	762	206	864	678	186

Note: Test for Equality of Means: (***) denotes significant at 1%.

Table 2. Description of process innovations in Spanish manufacturing firms

	1994			1998			2002		
	Total sample	Small firms	Large firms	Total sample	Small firms	Large firms	Total sample	Small firms	Large firms
Mean	0.3545	0.2834	0.5455	0.344	0.2992	0.5097	0.3032	0.2537	0.4839
Variance	0.2291	0.2034	0.249	0.2259	0.21	0.2511	0.2115	0.1896	0.2511
Test Eq Props		7.1081***			5.6427***			6.0499***	
No Obs	852	621	231	968	762	206	864	678	186

Note: Test for Equality of Proportions: (***) denotes significant at 1%.

Table 3. Description of share of skilled workers in Spanish manufacturing firms

	1994			1998			2002		
	Total sample	**Small firms**	**Large firms**	**Total sample**	**Small firms**	**Large firms**	**Total sample**	**Small firms**	**Large firms**
Mean	0.0821	0.0688	0.1179	0.0911	0.0800	0.1321	0.0972	0.0869	0.1346
Variance	0.0105	0.0088	0.0134	0.0118	0.0106	0.0142	0.0160	0.0156	0.0155
Test Eq means		5.7782***			5.7237***			4.6027***	
No Obs	852	621	231	968	762	206	864	678	186

Note: Test for Equality of Means: (***) denotes significant at 1%.

Table 4. TFP by innovation and human capital for small and large firms

		Small firms				Large firms			
		HK: 1st Q	HK: 2^{on} Q	HK: 3rd Q	HK: 4^{rt} Q	HK: 1st Q	HK: 2^{on} Q	HK: 3rd Q	HK: 4^{rt} Q
	1994	0.0164	0.0090	0.0442	0.0828	0.0375	0.0708	0.1378	0.1011
Innovative firms	1998	0.0539	0.0900	0.1279	0.1846	0.0949	0.1052	0.1711	0.2721
	2002	-0.0014	0.0952	0.1604	0.2005	0.1136	0.1739	0.1619	0.2465
	1994	-0.0716	-0.0175	-0.0120	0.0621	-0.0109	-0.0148	0.1199	0.1709
Non-innovative firms	1998	-0.0582	0.0390	0.0982	0.1036	0.0032	0.1503	0.1457	0.1325
	2002	0.0355	0.1064	0.1170	0.1427	0.0856	0.1755	0.1567	0.2173

Note: 1st, 2nd, 3rd and 4th HK are the observations corresponding to the different quartiles of human capital.

Table 5. Estimation results of the TFP specification

	OLS						RE		
	Baseline specification			**Additional control variables**			**Additional control variables**		
	(i) Total	**(ii) Small**	**(iii) Large**	**(iv) Total**	**(v) Small**	**(vi) Large**	**(vii) Total**	**(viii) Small**	**(ix) Large**
constant	0.0652**	0.0445	0.1652*	0.1261**	0.0888	0.2514**	0.0782	0.0644	0.2412
	(0.0291)	(0.0352)	(0.0949)	(0.0563)	(0.0649)	(0.1265)	(0.0713)	(0.0817)	(0.1518)
INN	0.0344***	0.0355***	0.0374***	0.0321***	0.0306***	0.0389***	0.0162**	0.0102	0.0355***
	(0.0092)	(0.0113)	(0.0158)	(0.0091)	(0.0112)	(0.0159)	(0.0080)	(0.0099)	(0.0129)
HK	0.2112***	0.1793***	0.3028***	0.1418***	0.1068**	0.2333***	0.0912**	0.0569	0.1626*
	(0.0458)	(0.0532)	(0.0897)	(0.0462)	(0.0543)	(0.0927)	(0.0424)	(0.0494)	(0.0865)
Size	0.0085**	0.0115*	0.0025	-0.0055	-0.0036	-0.0061	-0.0050	-0.0102	-0.0065
	(0.0035)	(0.0063)	(0.0133)	(0.0049)	(0.0075)	(0.0137)	(0.0053)	(0.0081)	(0.0156)
Age	0.0010***	0.0013***	0.0006*	0.0007***	0.0009***	0.0005	0.0006**	0.0005	0.0008**
	(0.0002)	(0.0003)	(0.0004)	(0.0002)	(0.0003)	(0.0004)	(0.0003)	(0.0004)	(0.0004)
$DSEC_t$	YES	YES	YES	YES	YES	YES	YES	YES	YES
$DYEAR_t$	YES	YES	YES	YES	YES	YES	YES	YES	YES
Control Vars									
Prodcap				0.0637**	0.06*	0.1115**	0.1159***	0.1224***	0.1166**
				(0.0323)	(0.0369)	(0.0585)	(0.0312)	(0.0360)	(0.0534)
Foreignk				0.0347**	0.0039	0.0515***	0.0506***	0.0391	0.0518***
				(0.0154)	(0.0263)	(0.0209)	(0.0154)	(0.0278)	(0.0190)
Group				0.0076	0.0306	-0.0158	-0.0101	0.0137	-0.0373**
				(0.0139)	(0.0199)	(0.0206)	(0.0130)	(0.0192)	(0.0184)
Publick				-0.0401	-0.0747	-0.0434	-0.0866*	-0.1546	-0.0897
				(0.0520)	(0.0964)	(0.0655)	(0.0486)	(0.0761)	(0.0625)
Export				0.0013**	0.0016**	0.0011	0.0014**	0.0016***	0.0010
				(0.0006)	(0.0007)	(0.0017)	(0.0006)	(0.0007)	(0.0013)
$DMKT_t$				YES	YES	YES	YES	YES	YES
$DREG_t$				YES	YES	YES	YES	YES	YES
RE_t							YES	YES	YES

(*continued*)

Table 5. *Continued*

H_0: $DSEC_t=0$	19.89***	14.44***	53.34***	16.12***	12.37***	9.12***	205.54***	156.51***	145.11***
H_0: $DYEAR_t=0$	38.72***	27.05***	13***	34.66***	26.11***	10.46***	115.77***	81.86***	30.59***
H_0: $DMKT_t=0$				2.23**	1.77	0.98	12.05**	10.8**	5.46
H_0: $DREG_t=0$				5.55***	5.16***	2.89***	62.54***	50.67***	33.98***
H_0: $RE_t=0$							571.58***	423.35***	58.99***
R^2 (adj)	0.2053	0.1868	0.2887	0.248	0.2361	0.3488	---	---	---
No Obs	2684	2061	623	2684	2061	623	2684	2061	623

Notes: Standard deviation based on robust estimation of the covariance matrix of coefficients in parentheses. DSEC=0 denotes an F-statistic to check the significance of industrial dummies, DYEAR=0 the significance of yearly dummies, DMKT=0 the significance of the dummies related to firm's main market and DREG=0 the significance of regional dummies. RE=0 is the Hausman test to select between the random and the fixed effect model.

(***), (**) and (*) denote significant at 1%, 5% and 10%.

Table 6. Decomposition of the TFP gap between large and small firms

	1994		1998		2002	
$\overline{TFP}_L - \overline{TFP}_S$	0.0793		0.0766		0.0602	
	Characteristics	**Returns**	**Characteristics**	**Returns**	**Characteristics**	**Returns**
PANEL A						
Total	0.0752	0.0041	0.0804	-0.0038	0.0690	-0.0087
	94.79%	5.21%	104.92%	-4.92%	114.49%	-14.49%
INN	0.0084	-0.0002	0.0068	-0.0001	0.0074	-0.0005
	10.60%	-0.03%	8.82%	-0.31%	12.27%	-0.81%
HK	0.0050	0.0121	0.0059	0.0142	0.0051	0.0146
	6.23%	15.25%	7.72%	18.50%	8.52%	24.30%
INN+HK	0.0134	0.0121	0.0127	0.0139	0.0125	0.0142
	16.84%	15.22%	16.54%	18.19%	20.78%	23.49%
PANEL B						
Total	0.0708	0.0111	0.0781	0.0196	0.0637	-0.0117
	89.28%	13.99%	101.96%	25.59%	105.81%	-19.43%
INN	0.0043	-0.0004	0.0034	-0.001	0.0037	-0.0018
	5.42%	-0.50%	4.44%	-1.30%	6.15%	-2.99%
HK	0.0032	0.01	0.0038	0.0116	0.0033	0.0121
	4.03%	12.61%	4.96%	15.14%	5.48%	20.10%
INN+HK	0.0075	0.0096	0.0072	0.0106	0.0070	0.0103
	9.45%	12.10%	9.40%	13.84%	11.63%	17.11%

Note: Given that the decomposition is not exact when using estimates from the RE model (Panel B),
 the sum of the shares of the components does not equal 100%.