Endogenous Productivity and Unobserved Prices*

Charlotte Guillard† Jordi Jaumandreu‡ Jocelyn Olivari§

January 29, 2018

Abstract

We discuss ways to apply the model for endogenous productivity when there are no firm-level output price indices available, a limitation of many data bases. Including the demand of the firm in the estimation allows us to obtain a "composite" of productivity, demand elasticity, and demand heterogeneity. This unobservable, often called "revenue productivity", is the estimate of productivity used by most scholarly studies. We find that this composite does not behave as productivity and, in particular, neither is greater for firms that perform R&D nor its distribution shows stochastic dominance. Its persistence and returns also give different results. Our findings highlight that results based on revenue productivity can be highly misleading about the returns of firm investments.

Keywords: R&D, productivity, production function.

*We are grateful to D. Aboal, U. Doraszelski, R. Alvarez, G. Carlomagno, G. Crespi, J.M. Gallego, J. Mairesse, P. Mohnen and Heng Yin, as well as numerous seminar audiences, for comments. This paper originated as an input for the IDB research project "Innovation Effort and Productivity Dynamics: Evidence for Latin America Countries." We thank Aptech for facilitating the use of Gauss.

†BETA, Strasbourg University and UNU-MERIT, email: charlotte.guillard@gmail.com.

‡Boston University, email: jordij@bu.edu.

§Universidad Adolfo Ibanez, email: jocelyn.olivari@uai.cl.

1. Introduction

Endogenous productivity results from the investments in R&D and innovation of firms, in imperfectly competitive markets, to reduce their costs and compite with better prices as well as to enhance directly the demand for their products. The explicit recognition that productivity is endogenous is important to measure it precisely and to assess the impact of these investments. Several recent papers have adopted this direction.[1] Important goals are assessing the impact of R&D and innovation expenditures, compare the resulting productivity of firms with and without these expenditures, and measure the returns. These assessments are decisive at the time to define policies.

This paper explores ways to estimate endogenous productivity using a rich data set, the 23 years of firm-level data generated by the Spanish ESEE survey (1990-2012). We replicate the estimation of the model in Doraszelski and Jaumandreu (2013), henceforth DJ, using the more than double years of data now available (the original model was estimated with the 10 years of the period 1990-1999). We compare the estimated coefficients, the results of the stochastic dominance tests, and the returns, verifying that they stay the same despite the additional 13 years included in the present paper. Then we briefly assess how important is the endogenenous treatment of inputs and productivity by comparing the results with the productivity measurements obtained using two traditional nonparametric approaches that treat them as exogenous: the productivity rates of growth of the Solow residual (Solow, 1957), and the productivity levels obtained with the Multilateral index of Caves, Christensen and Diewert (1982). DJ produces better production function estimates and more precise and contrsated productivity measures, that reflect the impact of the investments.

The insights of the DJ model are produced using firm-level output price indices that are available in the ESEE. However, many data bases have no information on output prices and only industry deflators are available. This is particularly true for many Latin American countries, and is the main motive of our paper. The estimates without firm-level output prices are prone to the criticisms raised by Klette and Griliches (1996). We discuss ways to apply the DJ model when there are no firm-level prices available. This is done by including in the estimation the demand function of the

[1] See, for example, Doraszelski and Jaumandreu (2013); Arx, Roberts and Xu (2011); De Loecker (2013); Peters, Roberts, Vuong and Fryges (2017); Boler, Moxnes and Ullveit-Moe (2015); Maican and Orth (2015) and Bilir and Morales (2016).

firm to make up for output prices, as Klette and Griliches (1996) suggested and De Loecker (2011) applied recently. It creates the possibility of estimation of a "composite" of productivity, demand elasticity, and demand heterogeneity, sometimes called "revenue productivity", that is being taken as an estimate of productivity in many recent works.[2]

This is, however, far away from being an ideal solution. On the one hand, it treats the composite as a Markov process although it is likely not to be such thing (if heterogeneity of production and demand function are characterized by distinct Markov processes, their sum cannot be a Markov process). On the other, it conveys in productivity three unobservables that are likely to be correlated among them, two of them distinct from productivity.[3] Estimation gives a prominent role to the estimation of demand elasticity since it nests the estimation of the production function in the demand function. Demand elasticity weights the sum of the unobservables productivity and demand heterogeneity.

This paper shows ways to estimate an endogenous productivity model without firm output prices despite all its problems. Then applies it to the data and analyzes the results. We look at the estimates of the coefficients of the production function, elasticity of demand, productivity and returns to R&D. The main conclusion that emerges is that the composite of productivity, unobserved demand heterogeneity and demand elasticity does not behave as a measure of productivity. To anticipate the main results let us say that the composite means are not systematically greater for R&D firms, and the distribution of the composite does not show stochastic dominance for the firms that invest in R&D. The implication is that researchers using this measure in applied trade, industrial organization or reallocation analysis, to cite a few areas, hardly can be reassured that they are getting results that describe the behavior of productivity.

The rest of this paper is organized as follows. Section two presents the data and Section three estimates productivity, tests stochastic dominance and estimates returns to R&D, with the exogenous indices. Section four is dedicated to the replication of the DJ exercise. Section five explains how to estimate the model without prices and its theoretical consequences, and Section six carries out its estimation and examines the results. Section seven is dedicated to compare the effects of R&D. Section eight concludes.

[2] Works that adopt the estimation of this composite are, for example, Hsieh and Klenow (2009); Gandhi, Rivers and Navarro (2013); Asker, Collard-Wesler and De Loecker (2014); Boler, Moxnes and Ullveit-Moe (2015); Peters, Roberts, Van Ahn and Fryges (2016) and Bilir and Morales (2016). On the contrary, Jaumandreu and Yin (2017) separate the unobservables with a method based on the availability of more than one market for each firm.

[3] Jaumandreu and Yin (2017) document the presence of a negative relationship between productivity and demand heterogeneity with a sample of Chinese firms.

2. Data

We use the Encuesta Sobre Estrategias Empresariales (ESEE) data corresponding to the period 1990-2012. It is a firm-level survey of Spanish manufacturing sponsored by the Ministry of Industry. At the beginning of the survey, about 5% of firms with up to 200 workers were sampled randomly by industry and size strata. All firms with more than 200 workers were included in the survey, and 70% of these larger firms responded. Firms disappear over time from the sample due to either exit (shutdown or abandonment of activity) or attrition. To preserve representativeness, samples of newly created firms were added to the initial sample almost every year. Over time, some additions counterbalanced attrition too.

We keep the firms for which we have enough information available and at least three consecutive years of observation. This gives a dataset with 3,026 firms and 26,977 observations. Detailed sample size and variable definitions can be found in the Data Appendix. We group the firms in the ten broadly defined industries also used in DJ.[4] Table A1 gives the industry definitions and Table A2 provides descriptive statistics, reporting means of the individual values in the industries.

There is an important peculiarity of the new data as compared with the one used in DJ. Firms' output fell sharply around 2008 in all industries, although with different intensity. Materials followed this output movement, but this was not the case for labor, and especially capital. Figure A1 depicts a few examples of this evolution in rates of growth. It reflects the behavior of firms' demand for inputs confronted to a sudden and unexpected fall in demand. It contrasts in deepness with the also acute previous recession in 1993. Recovery was complex and deserves further attention. From the point of view of the estimation of a production function and productivity, this is likely to create a significant underutilization of the inputs capital and labor that, if untreated, is likely to bias the estimated elasticities.

We have dealt with the event in two different ways. First we tried to employ the indicators of utilization of capacity, but we abandoned this method because of the likely endogeneity of this variable. Second, we have allowed for a change in the value of all the terms in capital and labor for the year 2008 by means of interacting yearly dummies. The value of the coeffients of these dummiesmay be interpreted as the estimated correction to be done to the observed input values to preserve the structural relationship.

[4] We add industry 5, absent in the original paper.

3. Estimating productivity with indices

In this section, we briefly comment on the application of two nonparametric measures of productivity: the Solow residual (rates of growth) and the Multilateral index (levels). The first measure is due to Solow (1957). Hall (1990) developed the way to apply Solow's idea with imperfect competition. The second measure is due to Caves, Christensen and Diewert (1982) who built an index in levels extending Solow's idea. Van Biesebroeck (2007 and 2008) summarizes these measures, among many other, and compares their performance. These measures are typically used in studies without focus on endogenous productivity.

Both measures have in common to be nonparametric estimates that approximate the input elasticities from observed input shares. In the absence of information on the economies of scale, it is usual to assume long run constant returns to scale. Then, assuming cost minimization, if one computes the input shares from total cost (as opposed to revenue), the Solow residual and the Multilateral index are robust to imperfect competition (Hall, 1990). The most usual method of calculation is to add an estimate of the cost of capital to variable costs.

Assume the production function is

$$Q_{jt} = F(K_{jt}, L_{jt}, M_{jt}) \exp(\omega_{jt} + e_{jt}), \tag{1}$$

where Q_{jt} is quantity produced, K_{jt}, L_{jt} and M_{jt} stand for the inputs capital, labor and materials respectively, ω_{jt} is persistent Hicks neutral productivity, and e_{jt} an error uncorrelated with all the information when the firm takes the decisions on output and inputs (e.g. an observational error). This is the general setting considered by the recent papers aimed at the estimation of the production function under the presence of persistent productivity.[5] These papers take capital as given and labor and materials as variable in the short run.

The Solow residual is

$$(\widehat{\omega_{jt} - \omega_{jt-1}})_S = (q_{jt} - q_{jt-1}) - (1 - \overline{S}_{Ljt} - \overline{S}_{Mjt})(k_{jt} - k_{jt-1}) - \overline{S}_{Ljt}(l_{jt} - l_{jt-1}) - \overline{S}_{Mjt}(m_{jt} - m_{jt-1}),$$

where the variables are in logs and the input shares are computed as

$$\overline{S}_{Xjt} = \frac{S_{Xjt} + S_{Xjt-1}}{2}, \qquad X = L, M.$$

[5] Olley and Pakes (1996); Levisohn and Petrin (2003); Ackerman, Caves and Frazer (2015). Gandhi, Rivers and Navarro (2013) can be considered a paper in this line but circumscribed itself to the estimation of production functions under perfect competition.

The multilateral index is

$$\widehat{(\omega_{jt}-\overline{\omega})}_M = (q_{jt}-\overline{q}) - (1-\tilde{S}_{Ljt}-\tilde{S}_{Mjt})(k_{jt}-\overline{k}) - \tilde{S}_{Ljt}(l_j-\overline{l}) - \tilde{S}_{Mjt}(m_{jt}-\overline{m}),$$

where the variables are in logs, and

$$\begin{aligned}\overline{x} &= \frac{1}{N}\sum_j \frac{1}{T_j}\sum_t \ln X_{jt}, \qquad X = K, L, M,\\ \tilde{S}_{Xjt} &= \frac{S_{Xjt}+\overline{S}_X}{2}, \qquad X = L, M,\end{aligned}$$

with N being the number of firms and T_j the number of time observations of firm j.

These measures have a main drawback: the estimation of the elasticities, in which the computations are based, is in general not consistent. Input shares in cost are independent of Hicks neutral productivity, but they are only a proper measure of elasticities if there are not adjustment costs (estimation also neglects e_{jt}). As capital is subject to high adjustment costs, and the same is likely to be the case for labor and materials (although adjustment costs of materials are generally considered smaller), the elasticities are likely to be biased in several directions. We do not know how these biases relate to productivity, and in particular endogenous productivity, and hence how the Solow residual and the Multilateral index going to perform with respect to the measurement of productivity. This is why these measurements have limited usefulness. But we can take them as a first approximation to have an idea of the main trends present in the data.

Tables 1 and 2 show the results of computing the Solow residual and the Multilateral index using the ESEE data base. We compute shares from total cost after adding an estimate of the cost of capital to variable costs. The user cost of capital is calculated from firm-level data on the interest rate paid for investments, an industry estimate of depreciation and the yearly rate of variation of the price of investment (see the Data Appendix). The tables show first the estimated input shares and then apply to the productivity measures the same tests that we will later apply to the parametric estimates. The tables show that shares are very similar, so we will focus on the productivity results.

Before comenting the results, let us make some remarks that are valid for all the tables of this paper in which we compute the tests. First, we compute the mean productivity for the subsample of observations with and without R&D, then we compute the statistic reported in page 1365 of DJ of the difference of means. A column gives the difference of means, another the value of the statistic and a third the probability value. The null hypothesis is that the mean of R&D is greater, and a probability for the value of the test below 5% is taken as the rule to reject it. We apply the test separately for the subsamples of firms with 200 workers or less and more than 200 workers, to respect the different representativity of the data for these two strata. Second, we compare each whole distribution (200 workers or less and more than 200 workers) by means of the Kolgomorov-Smirnov test of stochastic dominance reported in page 1367 of DJ. We first test the null hypothesis of equality of the distributions and then the null hypothesis that the distribution of the observations with R&D stochastically dominates. We apply the test using the mean productivity over time of each firm to avoid statistical dependence of the observations. Again, we split the sample for each industry in firms with 200 workers and less and firms with more than 200 workers. We only apply these tests when the subsamples of R&D and non R&D firms have at least 20 firms each.[6] Notice that Table 1 applies these tests to the Solow growth rates and Table 2 to the productivity levels computed with the Multilateral index.

Table 1 shows in columns (4) and (5) that average productivity growth is around 1% for the smaller firms and a little less than 0.5% for firms with 200 workers and more. The differences between the mean growth of productivity for firms with and without R&D are very small. Column (6) shows that average productivity growth is greater for firms with R&D in 11 cases out of 20. There is not discernible pattern across industries or sizes. But the average differences are so small in relation to the variance of productivity that in columns (7) and (8) is only possible to reject that the growth is higher for firms with R&D in 2 cases out of 20. The difficulty at the time of distinguishing the distributions of productivity becomes apparent in columns (9) and (10): the equality of the distributions of productivity growth can be rejected only in 2 cases out of 12. Given this result, the fact that stochastic dominance cannot be rejected in 10 out of 12 cases in columns (11) and (12) is unimportant. In only one out of these 10 cases the distributions are distinguishable.

In Table 2 the multilateral index gives in column (4) differences in the levels of productivity that range between -7.5% and 9% for the firms with and without R&D. Productivity is greater for R&D firms in 14 out of 20 cases, 7 out of 10 for the smallest firms and 7 out of 10 for the biggest. Here the negative mean differences are again small with repect to the variance of productivity, so that it is only possible to reject in 1 case out of the 20 that the firms with R&D do not have greater productivity. Columns (7) and (8) show that the equality of the distributions can be rejected in 5 cases out of 12. Consequently it is only partially informative that in columns (9) and (10) we cannot reject the stochastic dominance of the distribution of productivity for the firms with R&D in any industry. This is a much stronger statistical assessment if we have arrived to the conclusion that the

[6] This produces a total of 12 tests across industries against the 10 used in DJ because we have added industry 5 and we have now enough firms with more than 200 workers in industry 10.

distributions are distinct, and this happens only for 5 industries.

In summary, the indices show that firms with and without R&D seem to diverge in their levels of productivity, although this is imprecisely estimated. Average differences in productivity levels are below 9% and only fully significant in around half of the industries. But the Solow indice is not able to indicate any significant difference in the rates of growth. At this point it is hard to say if this is due to the absence of a greater growth of productivity for R&D firms or to the method applied to compute the growth. Later, we compare these results with the parametric results obtained controlling for endogeneity of the inputs and productivity.

4. Estimating endogenous productivity with output prices (DJ model)

We replicate the estimation of the model in DJ, using the 23 years of data now available. In this section, we briefly remind the model, report a few necessary small changes in the specification and describe the results.

DJ assumes a Cobb-Douglas production function and a Markov process for productivity ω_{jt} of the form $\omega_{jt} = \beta_t + g(\omega_{jt-1}, rd_{jt-1}) + \xi_{jt}$ where $rd_{jt-1} = \ln RD_{jt}$ is the log of expenditures in $R\&D$. Plugging this process in the production function (1), taking logs and using the inverted demand for labor[7] h_{jt-1} to replace ω_{jt-1}, we have

$$q_{jt} = \beta_K k_{jt} + \beta_L l_{jt} + \beta_M m_{jt} + \beta_t + g(h_{jt-1}, rd_{jt-1}) + \xi_{jt} + e_{jt}, \tag{2}$$

where we omit the constant for simlicity. This corresponds, with some slight and self-evident changes in notation, to equation (6) in DJ. Notice that we use an in-homogeneous Markov process, by specifying the time dummies β_t. We expect very little change, and this specification avoids to have 21 time dummies outside and inside the $g(\cdot)$ function with constrained coefficients that have to be treated as nonlinear parameters.

The inverted demand for labor is

$$\begin{aligned} h_{jt} = {} & \lambda - \beta_K k_{jt} + (1 - \beta_L - \beta_M) l_{jt} \\ & + (1 - \beta_M)(w_{jt} - p_{jt}) + \beta_M (p_{Mjt} - p_{jt}) - \ln\left(1 - \frac{1}{\eta(p_{jt}, d_{jt})}\right), \end{aligned} \tag{3}$$

where λ is a constant, w_{jt}, p_{Mjt} and p_{jt} represent wage, price of materials and output price, respectively, and d_{jt} stands for the firm-specific demand shifter. The demand shifter plays the role of a varying demand intercept. $\eta(\cdot)$ is a flexible function of the price and the shifter.

[7] The inverted demand for materials can also be used.

To allow the function $g(h_{jt-1}, r_{jt-1})$ to differ when the firm has chosen $R_{jt-1} = 0$, we specify it, as in DJ, as

$$1(RD_{jt-1} = 0)(g_{00} + g_{01}(h_{jt} - \lambda)) + 1(RD_{jt-1} > 0)(g_{10} + g_{11}(h_{jt} - \lambda, rd_{jt-1})).$$

We introduce two small changes in the specification. First, we reduce the terms of the polynomial that models $\eta(\cdot)$ in DJ from 9 to 6. To avoid excessive collinearity, we drop the terms featuring powers of d_{jt}. Second, we deal with the sharp variation in utilization of capacity in the year 2008 by including dummies interacted with the terms in capital and labor (see the Data section).

Estimation is by nonlinear GMM. We form the moments by using the constant, dummies and polynomials of the exogenous variables as instruments. The complete list is formed by the following instruments. First, the constant, 21 time dummies and the indicator of lagged R&D expenditures. Second, the 6 terms of the polynomial in lagged price and lagged demand shifter (all the terms of a complete polynomial of order three minus the square and cube of the demand shifter). Then we use k_{jt}, m_{jt-1}, a polynomial of order three in l_{jt-1}, and a polynomial of order three in the variables $k_{jt-1}, (w_{jt-1} - p_{jt-1})$ and $(p_{Mjt-1} - p_{jt-1})$, but often we drop some cross-terms and powers that exacerbate errors in variables.[8] We add interactions of the indicator of lagged R&D expenditures with variables k_{jt-1}, $m_{jt-1}, (w_{jt-1} - p_{jt-1})$ and $(p_{Mjt-1} - p_{jt-1})$, and the amount of the lagged R&D expenditures with k_{jt-1}, $m_{jt-1}, (w_{jt-1} - p_{jt-1})$, $(p_{Mjt-1} - p_{jt-1})$, p_{jt-1} and mdy_{jt-1}.[9] There are 46 parameters to estimate. With the dropping of some instruments according to industries the degrees of freedom range from 10 to 32. So we estimate with somewhat less instruments than in the original article.

Table 3 reports the results. Columns (1) to (3) report the estimates of the production function parameters. They are reasonable and precisely estimated.[10] All returns to scale (except industry 7) are, as expected, slightly below unity. The smallest values are reached for industries 3 and 6,

[8] We avoid the powers of k_{jt} and m_{jt-1} because they often create problems. This is why in some cases we drop the quadratic and cubic terms of the polynomial in l_{jt-1} and different cross terms and powers of the complete polynomial in $k_{jt-1}, (w_{jt-1} - p_{jt-1})$ and $(p_{Mjt-1} - p_{jt-1})$. Specifically, we only use l_{jt-1} in industries 3, 9 and 10, and we replace the polynomial of order three by a polynomial in k_{jt-1} and the terms $(w_{jt-1} - p_{jt-1})$ and $(p_{Mjt-1} - p_{jt-1})$ in industry 6, the term k_{jt-1} and polynomials in $(w_{jt-1} - p_{jt-1})$ and $(p_{Mjt-1} - p_{jt-1})$ in industries 7 and 8, and the terms k_{jt-1}, $(w_{jt-1} - p_{jt-1})$ and $(p_{Mjt-1} - p_{jt-1})$ in industry 10.

[9] But we drop the interaction of the indicator with k_{jt-1} in industries 7 and 8, and the whole interactions in industries 9 and 10. Similarly we drop the ineraction with the amount of the expenditures with k_{jt-1} in industry 1, and with k_{jt-1} and m_{jt-1} in industries 2,4,7 and 8.

[10] From here on we adopt the practice to report first stage GMM estimates. We only use the second stage to compute the specification test.

slightly below 0.9. The estimates for the elasticities are very close to the estimates of the original article for materials, but tend to be a little smaller for capital and a little greater for labor. We are not aware of any particular reason for this. The sharp changes in input utilization around 2008, already mentioned and difficult to model, seem to be reasonably controlled by the year-specific point dummies. The specification tests (columns (5) and (6)), presumably more difficult to be satisfied with a 23 years sample, are passed at the 5% significance level in all industries but 5, 8 and 9.

The elasticity of demand (column (4)) is estimated around an absolute value of two for all industries. This meets the threshold that we expect from theory (absolute elasticity above unity) but it is a somewhat low value that does not show a high variation. However, the elasticity of demand is residual in this model, and the estimates were already not very good in the original DJ. Probably, they can be improved trying a richer estimation of the function $\eta(\cdot)$ or, even better, using a separated markup equation to estimate it (see next section). It is however not clear how much could this improve the general results.

The productivity results are quite similar to the results in DJ. Figure 1 depicts the productivity distributions for firms without and with $R\&D$. Column (7) reports the difference in the mean productivity for the firms performing R&D and for the firms that have no R&D expenses. The differences are positive in 10 out of 20 cases, but many negative difference are small and/or subject to a high variance, so the null hypothesis according to which the mean of productivity of the firms performing $R\&D$ is greater can not be rejected in 16 out of the 20 cases.

Columns (10) and (11) show that we can reject the equality of the distributions of productivity in 8 out of 12 cases. Stochastic dominance of productivity of the firms with R&D is tested in columns (12) and (13), showing that it cannot be rejected in 10 out of the 12 cases. Notice that in 7 out of the 10 cases with stochastic dominance the distributions are significantly different, so these are the strongest results with respect to the role of R&D. In the original DJ stochastic dominance could not be rejected with distributions significantly different in only 5 cases ou of the total 10.

We now compare these results with the results of the application of the Multilateral index. The value of the estimated elasticities are clearly different from the approximations of the elasticities of the Multilateral index. In the estimation of endogenous productivity, the elasticity of capital is greater and the elasticity with respect to labor is lower (except in industry 7). This is usually interpreted as the result of controlling for the endogeneity of the inputs (see, e.g., Olley and Pakes, 1996). On the other hand, the difference of mean productivity for firms with R&D with respect to firms without R&D is positive in some fewer cases, but these differences are quite often estimated greater. The endogenus productivity estimation leads to statistical difference between the distributions of productivity for R&D and no R&D firms in more cases (8 out of 12 against 5 out of 12), and gives stochastic dominance of the productivity of firms with R&D conditional to the distribution being different in more cases (7 out of 12 against 5 out of 12). In addition, it is important to notice that the correlation of cases is not perfect. The intersection of these 7 and 5 industries is only 3 industries.

Summarizing, treating inputs and productivity as endogenous produces different input coefficients, discriminates more the distributions of productivity, and estimates larger differences in many cases in which mean productivity of the *R&D* firms is greater.

5. Estimating endogenous productivity without output prices

Output price plays two key roles in the above estimation. First, it allows specifying the production function directly since we can construct the quantity index Q_{jt}, obtained by deflating revenue. Second, as equation (3) makes clear, the output price is an argument of the inverse demand for labor through marginal revenue. In the absence of output price, we need to replace it by observable variables. In what follows we discuss how.

A model without prices

We need to specify the demand for the ouput of the firm. Let us write

$$Q_{jt} = \alpha_0 P_{jt}^{-\eta_{jt}} \exp(z_{jt}\alpha + \delta_{jt} + s_{jt}), \tag{4}$$

where α_0 is a constant, z_{jt} is a vector of observed demand shifters, δ_{jt} represents the effect of all the unobserved shifters, and s_{jt} is an observational error uncorrelated with the included variables. We are going to consider δ_{jt} as persistent as productivity ω_{jt}. This is the effect ususally know as "demand heterogeneity". Its likely importance was first pointed out by Foster, Haltiwanger and Syverson (2008). For a review of literature and an asessment see Jaumandreu and Yin (2017). Notice that we are specifying a varying elasticity of demand η_{jt}.This is important to make the model fullycomparable to DJ, an we later explain how to identify this varying elasticity. Multiplying equation (4) by P_{jt}, we get a relationship for revenue R_{jt}

$$R_{jt} = \alpha_0 P_{jt}^{-(\eta_{jt}-1)} \exp(z_{jt}\alpha + \delta_{jt} + s_{jt}).$$

Recall the production function of equation (1). To simplify notation, let us use the shorthand F_{jt} for $F(\cdot)$. The production function implies, under cost minimization, a dual marginal cost that we can write separating the part computable in terms of observable variables from unobservable productivity

$$\begin{aligned} MC_{jt} &= \frac{VC_{jt}}{(\beta_L+\beta_M)F_{jt}}\exp(-\omega_{jt})\exp(v_{jt}) \\ &= \overline{MC}(VC_{jt},K_{jt},L_{jt},M_{jt})\exp(-\omega_{jt})\exp(v_{jt}) \\ &= \overline{MC}_{jt}\exp(-\omega_{jt})\exp(v_{jt}), \end{aligned}$$

where v_{jt} is a disturbance that comes from possible observational problems in VC_{jt} (which, for example, may stem from the presence of adjustment costs not fully accounted for), and $\overline{MC}_{jt}$ is a short hand for the function $\overline{MC}(VC_{jt},K_{jt},L_{jt},M_{jt})$.

Profit maximization implies that $P_{jt}=\frac{\eta}{\eta-1}MC_{jt}$.[11] Replacing P_{jt} in the revenue relationship, using as marginal cost the expression introduced above, we get the equation

$$R_{jt}=\alpha_0(\frac{\eta_{jt}}{\eta_{jt}-1})^{-(\eta_{jt}-1)}\ \overline{MC}_{jt}^{-(\eta-1)}\exp(z_{jt}\alpha+(\eta-1)\omega_{jt}+\delta_{jt}-(\eta-1)v_{jt}+s_{jt}).$$

The replacement of output price induces the presence of the persistent unobserved variable $(\eta-1)\omega_{jt}+\delta_{jt}$, a composite of productivity, demand elasticity, and demand heterogeneity. Many recent papers have recognized the relevance of this composite unobservable.[12] If we want to estimate this relationship we need to control for this unobservable that we will call abbreviately $\tilde{\omega}_{jt}$. In logs, we can write the revenue equation as

$$r_{jt}=\ln\alpha_0+\varphi_{jt}-(\eta_{jt}-1)\overline{mc}_{jt}+z_{jt}\alpha+\tilde{\omega}_{jt}+u_{jt},$$

where $\varphi_{jt}=-(\eta_{jt}-1)\ln\frac{\eta_{jt}}{\eta_{jt}-1}$ and $u_{jt}=-(\eta_{jt}-1)v_{jt}+s_{jt}$. A constrained version of this equation has been used, for example, by Aw, Roberts and Xu (2011). By inverting the starting demand and multiplying by Q_{jt} it is possible to get another potentially usable relationship similar to the one employed by De Loecker (2011).[13]

Let us assume that $\tilde{\omega}_{jt}$ follows an (in-homogeneous) endogenous Markov process, i.e. $\tilde{\omega}_{jt}=\beta_t+g(\tilde{\omega}_{jt-1},rd_{jt-1})+\xi_{jt}$. This is an unsatisfactory assumption because, if the components follow

[11] This is the relevant relationship under static (non-dynamic) pricing. Virtually all the literature on productivity assumes static pricing. See Jaumandreu and Lin (2017) for a departure of this assumption.

[12] See the introduction.

[13] Inverting demand, we have

$$P_{jt}=(\alpha_0)^{\frac{1}{\eta_{jt}}}Q_{jt}^{-\frac{1}{\eta_{jt}}}\exp(\frac{z_{jt}\alpha+\delta_{jt}+s_{jt}}{\eta_{jt}}),$$

and multiplying by Q_{jt}

$$R_{jt}=(\alpha_0)^{\frac{1}{\eta_{jt}}}Q^{1-\frac{1}{\eta_{jt}}}\exp(\frac{z_{jt}\alpha+\delta_{jt}+s_{jt}}{\eta_{jt}}).$$

separately Markov processes, the sum of two Markov processes is not in general a Markov process. Let us keep this in mind when we analyze the results. We replace $\tilde{\omega}_{jt-1}$ by a function of observable variables $h(.)$, with shorthand h_{jt-1} (note that we slightly abuse of notation by writing the function without tilde for simplicity). The revenue equation becomes

$$r_{jt} = \ln \alpha_0 + \varphi_{jt} - (\eta_{jt} - 1)mc(vc_{jt}, k_{jt}, l_{jt}, m_{jt}) + z_{jt}\alpha + \beta_t + g(h_{jt-1}, rd_{jt-1}) + \xi_{jt} + u_{jt}. \tag{5}$$

This equation is an alternative to equation (2) where we can try to estimate, without using output prices, the parameters of the production function, the elasticity of demand and the composite unobservable.

To obtain the function $\tilde{\omega}_{jt} = h(.)$, we can invert either the FOC for materials or the FOC for labor. In both cases, we can use the corresponding FOC only or solve for the system of the two FOCs obtaining and inverting the demand for either materials or labor. Although this gives four possibilities, in practice we have found that the best results are obtained solving for the system and inverting the demand for labor. The reasons are likely to be that the inclusion of both input prices lessens a little the problems of multicollinearity and that the labor input is subject to less measurement problems. We have hence used

$$\begin{aligned} h_{jt} &= \lambda + \varphi_{jt}^{'} \\ &\quad + \eta(w_{jt} + l_{jt}) - (\eta - 1)\beta_K k_{jt} - (\eta - 1)\beta_M (w_{jt} - p_{Mjt}) - (\eta - 1)(\beta_L + \beta_M) l_{jt} - z_{jt}\alpha. \end{aligned} \tag{6}$$

where $\lambda = -\beta_M \ln \beta_L$ and $\varphi_{jt}^{'} = -(\eta_{jt} - 1)\beta_M \ln \beta_M - \eta_{jt} \ln \frac{\eta_{jt}}{\eta_{jt} - 1}$ (again there is some abuse of notation in writing λ) .

From equation (6), we can recover $h_{jt} - \lambda$, that estimates $\tilde{\omega}_{jt}$ up to a constant. Pluggin (6) into (5) allows to recover $g(h_{jt-1}, rd_{jt-1})$ up to a constant as well. The difference between h_{jt} and $g(h_{jt-1}, rd_{jt-1})$ is an estimate of ξ_{jt}.

Identifying η

Equations (4) and (5) show how important becomes the estimation of the elasticity of demand in the absence of ouput price. Estimates of a sort of equation (5) have been carried out, for example,

The revenue relationship is

$$R_{jt} = (\alpha_0)^{\frac{1}{\eta_{jt}}} F_{jt}^{1 - \frac{1}{\eta_{jt}}} \exp(\frac{z_{jt}\alpha + (\eta - 1)\omega_{jt} + \delta_{jt} + s_{jt}}{\eta_{jt}}).$$

The composite $(\eta_{jt} - 1)\omega_{jt} + \delta_{jt}$ shows up divided by η_{jt}. In practice, this equation seems to have identification problems.

in De Loecker (2011) and Peters, Roberts, Van Ahn and Fryges (2016). De Loecker (2011) estimates an industry constant η using the industry data (in opposition to firm-specific data). Peters, Roberts, Van Ahn and Fryges (2016) assume that $\beta_L + \beta_M = 1$ (constant short-run marginal cost) and use the average margin of the firms in an industry to estimate the elasticity of demand. Both solutions amount to substitute a given previous estimate for parameter η.

A more satisfactory procedure is to estimate simultaneously the elasticity of demand, as Jaumandreu and Yin (2017) do in another context. Using the expression $MC_{jt} = \frac{VC_{jt}}{(\beta_L+\beta_M)Q_{jt}} \exp(e_{jt} + v_{jt})$ and the optimal pricing rule $P_{jt} = \frac{\eta_{jt}}{\eta_{jt}-1} MC_{jt}$, we obtain

$$\frac{R_{jt}}{VC_{jt}} = \frac{1}{\beta_L + \beta_M} \frac{\eta_{jt}}{\eta_{jt} - 1} \exp(e_{jt} + v_{jt}), \tag{7}$$

an equation for the observed price-average cost margin ($\ln \frac{R_{jt}}{VC_{jt}} \simeq \frac{R_{jt}-VC_{jt}}{VC_{jt}} = \frac{P_{jt}-AVC^{obs}_{jt}}{AVC^{obs}_{jt}}$). The observed margin depends on two unobservable variables of interest: the short run elasticity of scale $(\beta_L + \beta_M)$ and the elasticity of demand η_{jt}. Equation (7) alone cannot identify them separately, but equations (7) and (5) can be arranged into the system

$$\begin{cases} r_{jt} - vc_{jt} = -\ln(\beta_L + \beta_M) + \ln \frac{\eta_{jt}}{\eta_{jt}-1} + e_{jt} + v_{jt} \\ r_{jt} = \varphi_{jt} + \beta_t - (\eta_{jt} - 1)\overline{mc}_{jt} + z_{jt}\alpha + g(h_{jt-1}, rd_{jt-1}) + \xi_{jt} + u_{jt}, \end{cases} \tag{8}$$

which contains the parameters β_L and β_M and the elasticity of demand in both equations (h_{jt-1} includes β_L, β_M and η_{jt-1}). In this paper we will identify these parameters in the following way. We will write and estimate the relevant part of the first equation of the system as $-\ln(\beta_L + \beta_M) + \ln \frac{\eta_{jt}}{\eta_{jt}-1} = \alpha_0 + x_{jt}\alpha_1 \equiv a_{jt}$, where x_{jt} are relevant variables to explain the margin α_0 and α_1 parameters. Then we will impose the values of the elasticity $\eta_{jt} = \frac{(\beta_L+\beta_M)\exp(a_{jt})}{(\beta_L+\beta_M)\exp(a_{jt})-1}$ as a nonlinear restriction linking parameters β_L, β_M and η_{jt} in the second equation.

6. Results from the model without prices

The first equation of the system is estimated by NLS, the second by nonlinear GMM. In the first equation estimation that follows we are going to use a unique variable x_{jt}, the state of the specific market of the firm or market dynamism. We first tried to model a_{jt} as a constant, but the data strongly reject this specification. The reason is that price-average cost margin varies intensely during the period, particularly since the economic crisis of 2008.

The specification requires to introduce in the second equation the relevant observed demand shifters. We include the age of the firms, measured in years, the total expenditure in advertising and (again) the indicator of market dynamism. Notice that we also use this variable at the time of modelling the elasticity of demand, so it can have two effects on demand: on its expansion and in the variation of its elasticity.

We form again the moments by using the constant, dummies and polynomials of the exogenous variables as instruments. The complete list is formed by the following instruments. First, the constant, 21 time dummies and the indicator of lagged R&D expenditures. Second we use a polynomial of order three in age_{jt} and lagged adv_{jt-1}, adding mdy_{jt}, and polynomials in the lagged price of the inputs wg_{jt-1} and p_{Mjt-1}. Then we use k_{jt} or k_{jt-1} together with a polynomial of order three in m_{jt-1} and l_{jt-1}.[14] We add a polynomial in lagged R&D expenditures and interactions of lagged R&D expenditures with variables k_{jt-1} and mdy_{jt}. In this case there are 44 parameters to estimate. With the dropping of some instruments according to industries the degrees of freedom range from 7 to 14.

Table 4 presents the results of estimating the model without firm-level output price. The estimates of the parameters of the production function are, in the case of labor and materials, very similar to the estimates of the DJ model. However, the coefficient on capital is here greater in all industies. The sum of the coefficients on labor and materials is almost always under but close to unity (except in three industries in which it sligthly exceeds one). But, with a greater coefficient on capital, this means that the long run elasticity of scale is here clearly above the estimate with DJ. This can be related to the imperfect control of the unobservable by means of a unique Markov process. There may be some persistent demand effects, positively correlated with capital, that determine some upward bias of its coeffient.

The model without prices estimates significantly greater elasticities of demand. As we now have a distribution of elasticities we report the median elasticity. Its value implies more realistic (lower) profitability rates that the estimates of the DJ model. Taking into account that the short-run rate of economic profitability can be written as $\frac{\pi_{jt}}{R_{jt}} = 1 - v + \frac{v}{\eta_{jt}}$, it is easy to see that it ranges from 9% to 15%. The specification test is in turn passed in all industries except 8, in which we are not able to estimate the second GMM stage.

Recall that what we get in this model is an estimate of the composite $\widetilde{\omega}_{jt} = (\eta_{jt} - 1)\omega_{jt} + \delta_{jt}$. Unlike the DJ model, this composite is not a measure of productivity, it is a mix of productivity

[14] We replace the polynomial in l_{jt-1} and m_{jt-1} by a polynomial of order three in each one of the variables in industries 5,6,8; by l_{jt-1} and a polynomial in m_{jt-1} in industry 10, and by l_{jt-1} and m_{jt-1} in industry 7. We use k_{jt} in indudtries 1,2,5,6,7 and 9; k_{jt-1} in industries 3,4,8, and both variables in industry 10. In 5 and 6 we use in fact a polynomial in the capital variable.

and unobserved demand heterogeneity weighted by the elasticity of demand. This makes it difficult to compare the results between the two models and explains the presence of important differences in the results. Notice that, taking the three involved variables η_{jt}, ω_{jt} and δ_{jt} as randon variables, the expectation of the composite unobservable is

$$E[(\eta - 1)\omega + \delta] = E(\eta - 1)E(\omega) + E(\delta) + Cov(\eta - 1, \omega).$$

To analyze the results, we scale the unobservable by $\overline{\eta} = mean(\eta_{jt})$, so the distributions that we draw and compare with the estimates of productivity obtained in DJ are $\frac{(\eta_{jt}-1)\omega_{jt}+\delta_{jt}}{\overline{\eta}}$. One can understand the results as showing an approximation to $E(\omega)$, plus the component $E(\delta)/\ \overline{\eta}$ added by the introduction of the demand relationship, plus the effect of the scaled covariance between productivity and elasticity of demand.

In the model without output prices, average productivity of the firms with R&D is greater than average productivity of the firms without R&D only in 4 cases out of 20 (against 10 cases out of 20 in the DJ estimation). On the other hand, now the equality of the distributions of the composite is rejected only in 4 out of 12 cases (against 8 cases out of 12 in the DJ estimation). In fact there is only 1 case (against 7 in the DJ estimation) in which the stochastic dominance can be established at the same time that the difference of the distributions. Hence, in addition of not being a measure of productivity, the composite does not behave as a measure of productivity.

To gain intuition about what can explain our results we first took a look at the distribution of the demand elasticity for firms with and without R&D. Both the shape of the distribution and average elasticity are, however, quite similar for both types of firms. This excludes the different elasticity of demand as the explanation for the frequent reversion of $\widetilde{\omega}_{jt}$ with respect to the value that is expected for a measure of productivity.

We should consider two types of possible biases. The first is related to the unsatisfactory asumption that the composite $\widetilde{\omega}_{jt}$ follows a Markov process. Our estimate of the composite may have the theoretical form stated above plus a bias induced by its wrong specification. The second, is the bias as a measure of ω_{jt}. The expectation of the scaled $\widetilde{\omega}_{jt}$, even estimated without bias, diverges from ω_{jt} for two motives: the presence of the (scaled) demand effect δ_{jt}, and the presence of a bias determined by the (scaled) covariance between elasticity of demand and productivity.

The estimates of δ_{jt} available elsewhere point to a negative correlation between ω_{jt} and δ_{jt} (Jaumandreu and Yin, 2017), with greater δ_{jt} for the R&D firms, so it would be a little surprising that this term is the responsible for the reversion of the stochastic dominance. A more likely reason is differences in the covariance term between R&D and no R&D firms. For example, firms reacted in Spain to the recession in domestic demand exporting more, and hence selling more in more competitive markets with higher elasticities of demand and lower margins. If this correlation is bigger for the set of no R&D firms, this would imply -according to the formula- a positive bias in the measurement of their productivity when we are not using firm-level output prices. In any case, the detailed study and test for all these biases is out of the scope of this paper and left for further research.

7. The effects of R&D.

Tables 5 and 6 summarize what models DJ and without firm-level output prices say, respectively, about the effects of R&D on productivity. We are going to discuss in turn the derivatives of ouput with respect to R&D and already attained productivity, productivity growth and the return to R&D.

The partial derivatives $\frac{\partial g(\omega_{jt-1}, rd_{jt-1})}{\partial rd_{jt-1}}$ and $\frac{\partial g(\omega_{jt-1}, rd_{jt-1})}{\partial \omega_{jt-1}}$ are, in both models, the elasticity of output (through expected productivity or expected composite unobservable) with respect to R&D expenditure and with respect to already attained productivity or composite unobservable. We have found that the statistics of these derivatives are not very sensitive to the trimming of extreme values, so we report them without any trimming.

Columns (1) to (4) of each table report the quartiles of the distribution of the elasticity with respect to R&D and a mean of these elasticities weighted by the sales of firms. The elasticities of output with respect to R&D in the DJ model are quite comparable to the results of the original article (although in three particular industries the estimated values are much greater). The presence of three slightly negative mean elasticities (only one in the original model) may be related to what has become recently clear in other papers. With differentiated products, more R&D can be associated to more quality of the goods, and hence less "net productivity" (productivity once quality has been deduced) as it is the productivity measured in this model (see Jaumandreu and Yin, 2017). In the case of the model with no firm-level output prices, the mean elasticity is negative in half of the cases and by an important amount. This confirms that the composite unobservable is not behaving as productivity.

Columns (5) to (10) of each table report the quartiles of the distribution of the elasticity with respect to attained productivity, splitted for R&D and no R&D firms. The elasticity of output with respect to already attained productivity in the DJ model diverges somewhat from the one in the original article. In contrast with the article, in which persistence of productivity (measured by a high derivative) was greater for the no R&D firms, now the pattern of the distributions are quite similar for performing and non-performing firms. Persistence has in fact tended to decrease for all firms, and more so for the no R&D firms. The most likely interpretation is that the 2000's, and in particular the crisis of 2008, sharply increased the uncertainty of productivity evolution, especially for the firms non performing R&D. The elasticity of output with respect to the composite of unobservables picked up by the model without firm-level output prices is very low, reflecting a dramatically smaller persistence for all kind of firms. This reflects that the heterogeneity of demand, and the own elasticity of demand, add an important source of random variability that lowers persistence.

Columns (11) to (13) show the expected productivity growth $g(\omega_{jt}, rd_{jt}) - g(\omega_{jt-1}, rd_{jt-1})$ as estimated by the models, total, and also split into R&D performers and non performers. In the case of the DJ model this is a real rate of productivity growth. In the case of the model without firm-level prices, our deflation of the composite unobservable by the consumer price index implies a sort of real rate as well.

The variation of expected productivity in the DJ model gives, for all firms, very reasonable rates of growth, a little greater than the productivity growth as measured by the Solow residual. Industry 2 shows no productivity growth, but this industry was already the one with the smallest growth when measured with the Solow residual. The growth of the expected part of the composite of unobservables picked up by the model without firm-level output prices is, for all firms and on average, greater. This is likely to reflect the evolution of the heterogeneity not included in the model for productivity.

When we compute the rates of growth for R&D and no R&D firms, two important facts emerge. First that productivity growth of the R&D firms is greater, in contrast with its level, only in three out of ten industries. Firms with R&D have greater productivity, but the growth of productivity is not necessarily greater. The second is that the growth of productivity according to the composite unobservable, despite we observe five cases of greater growth for R&D firms, seems not reliable for productivity. The measurements coincide in the ranking only in one industry.

Finally, in column (14) we report for both models net rates of return to R&D. We compute rates of return for each R&D performing firm as $[g(\omega_{jt}, rd_{jt}) - g(\omega_{jt-1}, rd_{jt-1})]\frac{V_j}{RD_j}$, where RD_j is the average over time of the firm expenditures in R&D and V_j the average over time of the firms value added. We use in the denominator an average of R&D expenditures to avoid a big volatily of the rates. As in DJ, we use value added to make the rates more comparable to classical estimates. To avoid the impact of the most extreme rates we only consider values between $\pm 2000\%$. We report the average weighted rates, using as a weight the R&D expenditures lagged two periods.

The rates of return computed with the DJ model are different, but basically comparable to the ones obtained with the original model. If the span of the data has more than doubled it seems sensible to assume that the rates have changed. The rates computed with the composite unobservable give some very high and some negative values, questioning the ability of this measure to provide any sensible indication about the returns to the investment in R&D.

8. Concluding remarks

We have explored ways to estimate endogenous productivity using the Spanish ESEE data base (1990-2012). First, we have replicated the estimation of the original model in DJ, using the additional 13 years of data now available, looking at the estimates and the results of the stochastic dominance tests. Results and tests are essentially the same as in the original article. Then, we have compared the results by applying two nonparametric measures (Solow residual and the Multilateral index), which take inputs and productivity as exogenous. The conclusion of this comparison is that treating inputs and productivity as endogenous produces better production function estimates and different, more discriminating, productivity measurements.

But many data bases do not have available firm-level prices, a key information used by DJ (and in the indices we have computed). In the absence of firm-level prices, we show that it is possible to estimate a compound of productivity and demand heterogeneity, weighted by the elasticity of demand. The estimation of this measure is however not straightforward. We need to estimate the elasticity of demand and the parameters of the production function, which introduces a significant problem of identification that has been only avoided until now with very restrictive solutions. Thus also implies to adopt the doubtful assumption according to which the unobservable composite follows a Markov process.

The composite of productivity and demand heterogeneity does not behave as productivity. It is in fact the result of two unobservables and an estimated weight, which may be correlated among them in non-obvious ways. In practice, the distribution of the composite for R&D firms doesn't show stochastic dominance over the distribution for no R&D firms. The elasticity of output with respect to R&D, the persistence of the composite, its growth and the returns to R&D give misleading values as well.

Data Appendix

We observe firms for a maximum of 23 years between 1990 and 2012. The sample is restricted to firms with at least three years of observations on all variables required for the estimation of the model. The number of firms with 3, 4,..., 23 years of data is 398, 298, 279, 278, 290, 324, 122, 111, 137,96, 110, 66, 66, 98, 66, 40, 37, 44, 37, 42 and 87 respectively. Table A1 gives the industry labels along with their definitions in terms of the ESEE, National Accounts, and ISIC classifications (columns (1)–(3)). Table A2 reports detail on the sample size and descritive statistics on the main variables.

Variable definitions are the same as in Doraszelski and Jaumandreu (2016), with the variables extended to the longer period of time 1990-2012:

- *Revenue* (R). Value of produced goods and services computed as sales plus the variation of inventories.

- *Price of output* (P). Firm-level price index for output. Firms are asked about the price changes they made during the year in up to five separate markets in which they operate. The price index is computed as a Paasche-type index of the responses.

- *Output* (Q). Value of produced goods and services computed as sales plus the variation of inventories deflated by a firm-specific price index of output.

- *Investment* (I). Value of current investments in equipment goods (excluding buildings, land, and financial assets) deflated by the price index of investment. The price of investment is the equipment goods component of the index of industry prices computed and published by the Spanish Ministry of Industry.

- *Capital* (K). Capital at current replacement values $\widetilde{K}_{jt}$ is computed recursively from an initial estimate and the data on current investments in equipment goods $\widetilde{I}_{jt}$. We update the value of the past stock of capital by means of the price index of investment P_{It} as $\widetilde{K}_{jt} = (1 - d)\frac{P_{It}}{P_{It-1}}\widetilde{K}_{jt-1} + \widetilde{I}_{jt-1}$, where d is an industry-specific estimate of the rate of depreciation. Capital in real terms is obtained by deflating capital at current replacement values by the price index of investment as $K_{jt} = \frac{\widetilde{K}_{jt}}{P_{It}}$.

- *Labor* (L). Total hours worked computed as the number of workers times the average hours per worker, where the latter is computed as normal hours plus average overtime minus average working time lost at the workplace.

- *Materials* (M). Value of intermediate goods consumption (including raw materials, components, energy, and services) deflated by a firm-specific price index of materials.

- *Variable cost* (VC). Wage bill plus the cost of materials minus all the expenditures destinated to advertising and R&D.

- *Wage* (W). Hourly wage cost computed as total labor cost including social security payments divided by total hours worked.

- *Price of materials* ($\mathrm{P_M}$). Firm-specific price index for intermediate consumption. Firms are asked about the price changes that occurred during the year for raw materials, components, energy, and services. The price index is computed as a Paasche-type index of the responses.

- *User cost of capital* ($\mathrm{P_K}$). Computed as $P_{It}(r_{jt} + \delta - CPI_t)$, where P_{It} is the price index of investment, r_{jt} is a firm-specific interest rate, δ is an industry-specific estimate of the rate of depreciation, and CPI_t is the rate of inflation as measured by the consumer price index.

- *Advertising* (Adv). Total expenditure in advertising.

- *R&D expenditures* (R&D). R&D expenditures include the cost of intramural R&D activities, payments for outside R&D contracts with laboratories and research centers, and payments for imported technology in the form of patent licensing or technical assistance, with the various expenditures defined according to the OECD Frascati and Oslo manuals.

- *Market dynamism* (mdy). Firms are asked to assess the current and future situation of the main market in which they operate. The demand shifter codes the responses as 0, 0.5, and 1 for slump, stability, and expansion, respectively.

- *Age* (age). Years elapsed since the foundation of the firm with a maximum of 40 years.

- *Firm size* (size). Number of workers in the year the firm enters the sample.

References

Ackerberg, D., K. Caves and G. Frazer (2015), "Structural Identification of Production Functions," *Econometrica*, 6, 2411-2451.

Aw, B.Y., M. Roberts and D.Y. Xu (2011), "R&D Investment, Exporting and Productivity dynamics," *American Economic Review*, 101, 1312-1344.

Asker, J., A. Collard-Wexler and J. De Loecker (2014), "Dynamic Inputs and Resource (Mis)Allocation," *Journal of Political Economy*, 22, 1013-1063.

Bilir, L.K. and E. Morales (2016), "The Impact of Innovation in the Multinational Firm," mimeo, University of Wisconsin-Madison and Princeton University.

Boler, E., A. Moxnes and K. Ullveit-Moe (2015), "R&D, International Sourcing, and the Joint Impact on Firm Performance," *American Economic Review,* 105, 3704-3739.

Caves, D.W., L.R. Christensen, and W. E. Diewert (1982), "Multilateral comparisons of output, input, and productivity using superlative index numbers," *The Economic Journal*, 92, 73-86.

De Loecker, J. (2011), "Product Differentiation, Multiproduct Firms, and Estimating the Impact of Trade Liberalization on Productivity," *Econometrica*, 79, 1407-1451.

De Loecker, J. (2013), "Detecting Learning by Exporting," *American Economic Journal: Microeconomics*, 5, 1-21.

Doraszelski, U. and J. Jaumandreu (2013), "R&D and Productivity: Estimating Endogenous Productivity," *Review of Economic Studies,* 80, 1338-1383.

Doraszelski, U. and J. Jaumandreu (2016), "Measuring the Bias of Technological Change," forthcoming in *Journal of Political Economy.*

Foster, L., J. Haltiwanger and C. Syverson (2008), "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?," *American Economic Review*, 98, 394-495.

Gandhi, A., S. Navarro and D. Rivers (2013), "On the Identification of Production Functions: How Heterogeneous is Productivity?," mimeo, University of Wisconsin-Madison.

Hall, R.E. (1990), "Invariance Properties of Solow's Productivity Measure,"
in Growth/Productivity/Unemployment: Essays to Celebrate Bob Solow's Birthday, Diamond, P. ed., MIT Press.

Hsieh, C.T. and P. Klenow (2009), "Misallocation and manufacturing TFP in China and India," *Quarterly Journal of Economics*, 124, 1403-1448.

Jaumandreu, J. and S. Lin (2017), "Pricing Under Innovation: Evidence from Spanish Manufacturing Firms," mimeo, Boston University and Xiamen University.

Jaumandreu, J. and H. Yin (2016), "Cost and Product Advantages: A Firm-level Model for the Chinese Exports and Industry Growth," Mimeo, Boston University.

Klette, T. and Z. Griliches (1996), " The Inconsistency of the Common Scale Estimators when Output Prices are Unobserved and Endogenous," *Journal of Applied Econometrics*, 11, 343-361.

Levinsohn, J. and A. Petrin (2003), "Estimating Production Functions Using Inputs to Control for Unobservables," *Review of Economic Studies*, 70, 317-341.

Maican, F. and M. Orths (2017), "Productivity Dynamics and the Role of 'Big-Box' Entrants in Retailing," *Jornal of Industrial Economics,* 65*,* 397-438.

Olley, S. and A. Pakes (1996), "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, 64, 1263-1298.

Peters, B., M. Roberts, V.A. Vuong and H. Fryges (2016), "Estimating Dynamic R&D Choice: An Analysis of Cost and Long-Run Benefits," *Rand Journal of Economics,* 48, 409-437..

Solow, R.M. (1957), "Technical Change and the Aggregate Production Function," *The Review of Economics and Statistics*, 312-320.

Van Biesebroeck, J. (2007), "Robustness of Productivity Estimates," *The Journal of Industrial Economics*, 55, 529-569.

Van Biesebroeck, J. (2008), "The Sensitivity of Productivity Estimates: Revisiting Three Important Debates," *Journal of Business & Economic Statistics*, 26, 311-328.

Table 1. Solow residual. Testing productivity growth of R&D and no R&D firms.

Industry	Shares[a] s_K (s. d.)	Shares[a] s_L (s.d.)	Shares[a] s_M (s. d.)	Prod. growth ≤ 200 (s.d.)	Prod. growth >200 (s.d.)	Diff. of means (≤ 200) (> 200)	Mean with R&D is greater t	Mean with R&D is greater p val.	Kolgomorov-Smirnov tests[b] Distributions are equal S_1	Kolgomorov-Smirnov tests[b] Distributions are equal p val.	Kolgomorov-Smirnov tests[b] Distribution with R&D dominates S_2	Kolgomorov-Smirnov tests[b] Distribution with R&D dominates p val.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
1. Metals and metal products	0.037	0.303	0.660	0.010	0.006	0.007	-0.990	0.839	1.398	0.040	0.602	0.484
	(0.026)	(0.145)	(0.153)	(0.133)	(0.067)	-0.001	0.084	0.467				
2. Non-metallic minerals	0.056	0.300	0.645	0.006	0.001	0.010	-0.851	0.802	0.842	0.478	0.140	0.961
	(0.038)	(0.138)	(0.145)	(0.144)	(0.062)	-0.021	2.271	0.012				
3. Chemical products	0.036	0.243	0.721	0.010	0.009	0.002	-0.219	0.587	0.749	0.629	0.749	0.326
	(0.024)	(0.122)	(0.129)	(0.107)	(0.077)	0.001	-0.078	0.531				
4. Agric. and ind. machinery	0.029	0.334	0.637	0.011	0.004	0.012	-1.204	0.885	0.940	0.340	0.940	0.171
	(0.021)	(0.140)	(0.147)	(0.140)	(0.065)	0.012	-0.781	0.782				
5. Electrical goods	0.032	0.307	0.661	0.011	0.005	-0.007	0.831	0.203	0.840	0.481	0.307	0.828
	(0.022)	(0.142)	(0.146)	(0.128)	(0.079)	-0.007	0.490	0.312				
6. Transport equipment	0.040	0.286	0.674	0.009	0.008	0.006	-0.554	0.710	0.828	0.500	0.774	0.302
	(0.026)	(0.151)	(0.160)	(0.104)	(0.089)	-0.006	0.513	0.304				
7. Food, drink and tobacco	0.037	0.232	0.731	0.008	0.001	-0.002	0.278	0.390	1.749	0.004	1.749	0.002
	(0.025)	(0.148)	(0.157)	(0.110)	(0.077)	0.003	-0.421	0.663	0.736	0.650	0.233	0.897
8. Textile, leather and shoes	0.032	0.346	0.622	0.009	0.002	0.000	-0.005	0.502	0.519	0.950	0.369	0.761
	(0.024)	(0.211)	(0.216)	(0.157)	(0.047)	-0.016	1.832	0.034	1.283	0.075	1.283	0.037
9. Timber and furniture	0.036	0.282	0.682	0.015	0.003	-0.015	0.925	0.180				
	(0.024)	(0.143)	(0.148)	(0.151)	(0.049)	-0.034	1.382	0.087				
10. Paper and printing products	0.050	0.300	0.650	0.013	0.003	0.006	-0.398	0.654	1.247	0.089	0.846	0.239
	(0.033)	(0.132)	(0.142)	(0.116)	(0.068)	0.010	-0.873	0.808	1.242	0.091	0.587	0.503
All industries	0.037	0.294	0.669	0.010	0.004							
	(0.027)	(0.155)	(0.162)	(0.130)	(0.068)							

[a] Input shares are estimated from total cost. Imperfect competition and constant returns to scale are assumed.
[b] Applied to a firm's average expected productivity when each sample has more than 20 firms.

Table 2. Multilateral index. Testing productivity levels of R&D and no R&D firms.

Industry	Shares[a] s_K (s. d.)	Shares[a] s_L (s.d.)	Shares[a] s_M (s. d.)	Diff. of means (≤ 200) (> 200)	Mean with R&D is greater t	Mean with R&D is greater p val.	Kolgomorov-Smirnov tests[b] Distributions are equal S_1	Kolgomorov-Smirnov tests[b] Distributions are equal p val.	Kolgomorov-Smirnov tests[b] Distribution with R&D dominates S_2	Kolgomorov-Smirnov tests[b] Distribution with R&D dominates p val.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1. Metals and metal products	0.037	0.315	0.648	0.087	-6.661	1.000	1.292	0.071	0.045	0.996
	(0.013)	(0.073)	(0.077)	0.023	-1.051	0.853				
2. Non-metallic minerals	0.054	0.302	0.644	0.046	-2.255	0.987	1.508	0.021	0.051	0.995
	(0.020)	(0.073)	(0.078)	-0.007	0.278	0.391				
3. Chemical products	0.035	0.249	0.716	0.077	-6.627	1.000	1.586	0.013	0.000	1.000
	(0.012)	(0.062)	(0.066)	0.040	-1.539	0.937				
4. Agric. and ind. machinery	0.028	0.342	0.630	0.071	-5.038	1.000	1.112	0.169	0.418	0.705
	(0.011)	(0.073)	(0.077)	-0.029	1.005	0.159				
5. Electrical goods	0.031	0.320	0.649	-0.002	0.073	0.471	1.595	0.012	0.186	0.933
	(0.011)	(0.071)	(0.074)	0.048	-1.999	0.976				
6. Transport equipment	0.039	0.299	0.661	0.067	-3.958	1.000	1.177	0.125	0.358	0.774
	(0.014)	(0.078)	(0.083)	0.024	-1.490	0.931				
7. Food, drink and tobacco	0.036	0.244	0.719	0.079	-5.558	1.000	2.135	0.000	0.314	0.821
	(0.013)	(0.076)	(0.081)	0.018	-1.200	0.884	0.813	0.524	0.364	0.767
8. Textile, leather and shoes	0.031	0.351	0.617	0.084	-7.317	1.000	1.943	0.001	0.056	0.994
	(0.012)	(0.106)	(0.109)	0.055	-3.078	0.999	1.129	0.156	0.095	0.982
9. Timber and furniture	0.036	0.286	0.678	-0.022	0.866	0.195				
	(0.013)	(0.072)	(0.075)	-0.074	1.577	0.061				
10. Paper and printing products	0.049	0.308	0.643	-0.070	3.315	0.001	1.069	0.203	1.069	0.102
	(0.017)	(0.067)	(0.072)	0.034	-1.396	0.918	0.863	0.446	0.035	0.998
All industries	0.036	0.303	0.661							
	(0.014)	(0.078)	(0.082)							

[a] Input shares are estimated from total cost. Imperfect competition and constant returns to scale are assumed.
[b] Applied to a firm's average expected productivity when each sample has more than 20 firms.

Table 3. Estimating and testing productivity: DJ model of endogenous productivity.

Industry	GMM[a] β_K (s. e.)	β_L (s.e.)	β_M (s. e.)	η (s.e)	χ^2 (df)	p val.	Diff. of means (≤ 200) (> 200)	Mean with R&D is greater t	p val.	Kolgomorov-Smirnov tests[b] Distrib. are equal S_1	p val.	Distrib. with $R\&D$ dominates S_2	p val.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
1. Metals and metal products	0.066	0.229	0.699	1.978	27.656	0.376	-0.001	0.023	0.491	1.460	0.028	0.837	0.246
	(0.033)	(0.067)	(0.015)	(0.081)	(26)		-0.011	0.392	0.348				
2. Non-metallic minerals	0.058	0.244	0.688	2.007	31.117	0.120	0.040	-1.515	0.934	1.653	0.008	0.131	0.966
	(0.058)	(0.075)	(0.020)	(0.227)	(23)		-0.043	1.548	0.061				
3. Chemical products	0.066	0.109	0.711	2.019	30.215	0.455	0.202	-7.021	1.000	2.305	0.000	0.121	0.971
	(0.041)	(0.109)	(0.031)	(0.058)	(30)		0.118	-2.674	0.996				
4. Agric. and ind. machinery	0.100	0.167	0.654	1.880	28.244	0.207	0.110	-6.842	1.000	1.831	0.002	0.706	0.369
	(0.039)	(0.095)	(0.037)	(0.059)	(23)		-0.006	0184	0.427				
5. Electrical goods	0.085	0.278	0.631	1.949	52.633	0.012	-0.029	1.998	0.023	0.608	0.853	0.357	0.775
	(0.039)	(0.095)	(0.037)	(0.095)	(32)		0.033	-1.385	0.916				
6. Transport equipment	0.063	0.147	0.669	1.998	19.283	0.375	0.116	-5.571	1.000	2.009	0.001	0.085	0.986
	(0.032)	(0.050)	(0.026)	(0.114)	(18)		0.001	-0.048	0.519				
7. Food, drink and tobacco	0.098	0.279	0.761	2.044	10.937	0.362	-0.037	0.711	0.239	3.007	0.000	3.007	0.000
	(0.044)	(0.094)	(0.010)	(0.087)	(10)		-0.087	4.302	0.000	1.306	0.066	1.306	0.033
8. Textile, leather and shoes	0.056	0.301	0.553	2.136	23.455	0.009	0.071	-5.965	1.000	1.813	0.003	0.089	0.984
	(0.031)	(0.054)	(0.016)	(0.537)	(10)		0.089	-4.907	1.000	1.370	0.047	0.133	0.965
9. Timber and furniture	0.050	0.251	0.631	1.929	65.250	0.000	0.035	-1.063	0.854				
	(0.047)	(0.076)	(0.060)	(0.174)	(26)		-0.117	1.877	0.035				
10. Paper and printing products	0.151	0.175	0.655	2.088	13.531	0.195	-0.102	4.611	0.000	0.980	0.292	0.980	0.147
	(0.027)	(0.053)	(0.021)	(0.195)	(10)		-0.024	0.943	0.173	1.001	0.269	1.001	0.135

[a] Reported coefficients are first stage estimates.
[b] Applied to the firm's average expected productivity when each sample has more than 20 firms.

Table 4. Estimating and testing productivity: Model without firm-level output prices.

Industry	GMM[a] β_K (s. e.)	GMM[a] β_L (s.e.)	GMM[a] β_M (s. e.)	GMM[a] η	GMM[a] χ^2 (*df*)	GMM[a] p val.	Diff. of means (≤ 200) (> 200)	Mean with R&D is greater t	Mean with R&D is greater p val.	Kolgomorov-Smirnov tests[b] Distrib. are equal S_1	Kolgomorov-Smirnov tests[b] Distrib. are equal p val.	Kolgomorov-Smirnov tests[b] Distrib. with *R&D* dominates S_2	Kolgomorov-Smirnov tests[b] Distrib. with *R&D* dominates p val.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
1. Metals and metal products	0.176	0.225	0.745	11.864	17.978	0.236	-0.047	0.721	0.236	0.333	1.000	0.333	0.801
	(0.026)	(0.025)	(0.034)		(14)		-0.090	2.158	0.016				
2. Non-metallic minerals	0.160	0.209	0.776	7.587	8.718	0.849	-0.003	0.038	0.485	0.600	0.864	0.600	0.487
	(0.039)	(0.037)	(0.052)		(14)		-0.088	0.781	0.218				
3. Chemical products	0.174	0.270	0.729	7.631	10.998	0.686	-0.102	0.955	0.171	1.012	0.258	0.585	0.505
	(0.042)	(0.060)	(0.125)		(14)		-0.016	0.224	0.411				
4. Agric. and ind. machinery	0.147	0.269	0.680	16.019	9.097	0.825	-0.026	0.223	0.412	1.152	0.140	1.152	0.070
	(0.024)	(0.027)	(0.032)		(14)		-0.030	1.059	0.145				
5. Electrical goods	0.188	0.292	0.624	24.546	4.325	0.987	0.048	-0.592	0.722	1.201	0.112	0.076	0.989
	(0.027)	(0.062)	(0.062)		(13)		-0.091	1.584	0.057				
6. Transport equipment	0.135	0.282	0.770	6.473	5.326	0.967	-0.045	0.677	0.250	0.837	0.485	0.289	0.846
	(0.027)	(0.062)	(0.161)		(13)		0.030	-0.570	0.716				
7. Food, drink and tobacco	0.152	0.083	0.924	7.143	1.974	0.961	-0.210	4.982	0.000	1.713	0.006	1.713	0.003
	(0.073)	(0.022)	(0.200)		(7)		-0.157	4.183	0.000	2.813	0.000	2.813	0.000
8. Textile, leather and shoes	0.132	0.393	0.549	20.951	-	-	-0.268	2.038	0.021	0.933	0.349	0.933	0.175
	(0.055)	(0.061)	(0.054)		(-)		-0.032	0.255	0.399	1.985	0.001	1.985	0.000
9. Timber and furniture	0.359	0.197	0.824	8.715	14.177	0.437	0.040	-0.473	0.681				
	(0.073)	(0.028)	(0.071)		(14)		-0.342	5.065	0.000				
10. Paper and printing products	0.200	0.203	0.738	12.555	6.106	0.806	0.088	-0.930	0.823	1.432	0.033	0.686	0.390
	(0.026)	(0.027)	(0.036)		(10)		-0.181	5.703	0.000	1.063	0.208	1.063	0.104

[a] Reported coefficients are first stage estimates.
[b] Applied to the firm's average expected productivity when each sample has more than 20 firms.

Table 5. Elasticities of output with respect to R&D and already attained productivity, productivity growth and rate of return to R&D (DJ model).

	Elasticity of output wrt. RD_{jt-1}				Elasticity of output wrt. ω_{jt-1}						Productivity growth			Net rate of return to R&D
					Performers			Non-performers						
Industry	Q1	Q2	Q3	Mean	Q1	Q2	Q3	Q1	Q2	Q3	Total	R&D	No R&D	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
1. Metals and metal products	-0.072	-0.043	0.015	0.146	0.438	0.543	0.645	0.463	0.474	0.479	0.049	0.036	0.104	0.368
2. Non-metallic minerals	-0.009	0.000	0.005	-0.016	0.449	0.501	0.557	0.395	0.465	0.520	-0.002	-0.007	0.013	0.198
3. Chemical products	-0.003	0.002	0.005	0.012	0.464	0.497	0.522	0.416	0.440	0.466	0.017	0.019	-0.003	0.683
4. Agric. and ind. machinery	-0.034	-0.012	0.009	-0.010	0.354	0.510	0.697	0.324	0.451	0.569	0.014	0.011	0.023	0.922
5. Electrical goods	-0.005	0.002	0.009	0.019	0.368	0.404	0.428	0.398	0.440	0.479	0.023	0.024	0.010	0.682
6. Transport equipment	-0.120	-0.039	0.122	0.127	0.381	0.697	0.878	0.406	0.448	0.470	0.020	0.017	0.026	0.211
7. Food, drink and tobacco	-0.027	0.000	0.010	-0.002	0.273	0.332	0.441	0.493	0.515	0.536	0.008	0.000	0.023	0.232
8. Textile, leather and shoes	0.026	0.0079	0.102	0.030	0.177	0.228	0.303	0.277	0.290	0.296	0.008	-0.003	0.014	0.161
9. Timber and furniture	-0.050	-0.025	0.053	0.082	0.422	0.504	0.647	0.375	0.419	0.484	0.016	0.035	0.009	0.522
10. Paper and printing products	-0.032	-0.017	0.030	0.031	0.230	0.320	0.472	0.289	0.336	0.382	0.010	0.005	0.015	0.454

Table 6. Elasticities of output with respect to R&D and already attained productivity, productivity growth and rate of return to R&D (model without firm-level output prices)

	Elasticity of output wrt. R_{jt-1}				Elasticity of output wrt. ω_{jt-1}						Productivity growth			Net rate of return
					Performers			Non-performers						
Industry	Q1	Q2	Q3	Mean	Q1	Q2	Q3	Q1	Q2	Q3	Total	R&D	No R&D	to R&D
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
1. Metals and metal products	-0.090	0.058	0.137	-0.062	0.123	0.203	0.250	0.087	0.087	0.088	0.039	0.056	-0.060	1.298
2. Non-metallic minerals	-0.131	-0.066	0.094	0.074	0.119	0.184	0.244	0.082	0.103	0.125	0.015	0.015	0.013	-0.048
3. Chemical products	-0.282	-0.125	0.173	0.199	0.116	0.278	0.505	0.265	0.394	0.470	0.114	0.119	0.068	-0.143
4. Agric. and ind. machinery	-0.074	-0.029	0.055	-0.026	0.020	0.072	0.158	0.117	0.150	0.165	0.015	0.006	0.043	0.519
5. Electrical goods	-0.059	0.032	0.103	-0.103	0.015	0.032	0.055	0.021	0.023	0.029	-0.012	-0.012	-0.011	0.883
6. Transport equipment	-0.063	-0.011	0.054	0.048	0.214	0.330	0.428	0.302	0.395	0.443	0.016	0.019	0.000	0.298
7. Food, drink and tobacco	-0.151	-0.036	0.086	0.009	0.093	0.135	0.238	0.110	0.127	0.158	0.030	0.040	0.016	0.407
8. Textile, leather and shoes	-0.472	-0.067	0.373	-0.101	0.078	0.125	0.192	0.030	0.090	0.215	0.010	0.023	0.033	-0.158
9. Timber and furniture	-0.108	-0.024	0.060	-0.041	0.101	0.141	0.170	0.107	0.119	0.127	0.046	0.037	0.060	0.914
10. Paper and printing products	-0.022	-0.010	0.009	0.002	0.024	0.058	0.122	0.049	0.090	0.142	0.041	0.002	0.079	0.983

Table A1. Industry definitions and equivalences.

	Industry	ESEE	National Accounts	ISIC (Rev. 4)
		(1)	(2)	(3)
1	Ferrous and non-ferrous metals and metal products	12+13	DJ	C 24+25
2	Non-metallic minerals	11	DI	C 23
3	Chemical products	9+10	DG-DH	C 20+21+22
4	Agricultural and industrial machinery	14	DK	C 28
5	Electrical goods	15+16	DL	C 26+27
6	Transport equipment	17+18	DM	C 29+30
7	Food, drink and tobacco	1+2+3	DA	C 10+11+12
8	Textile, leather and shoes	4+5	DB-DC	C 13+14+15
9	Timber and furniture	6+19	DD-DN38	C 16+31
10	Paper and printing products	7+8	DE	C 17+18

Table A2
Descriptive statistics

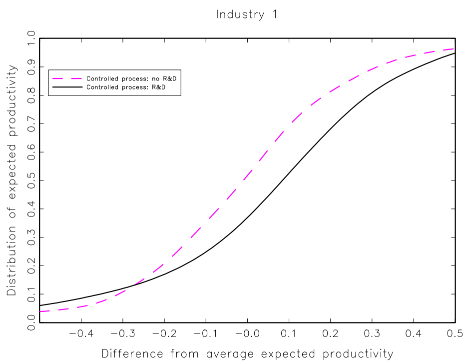
					Rates of growth								
Industry	Firms	Obs.	Entry (%)	Exit (%)	Revenue (s.d)	Price (s.d.)	Output (s.d.)	Labor (s.d.)	Capital (s.d.)	Materials (s.d.)	Variable cost (s.d)	Wage (s.d)	Price of mats. (s.d)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
1. Metals and metal products	433	3537	48.8	11.6	0.023	0.015	0.008	0.040	-0.010	-0.012	0.032	0.043	0.042
					(0.273)	(0.061)	(0.267)	(0.191)	(0.174)	(0.349)	(0.265)	(0.153)	(0.074)
2. Non-metallic minerals	228	1788	36.0	17.4	-0.006	0.011	-0.017	0.040	-0.027	-0.028	0.008	0.043	0.033
					(0.272)	(0.056)	(0.271)	(0.202)	(0.195)	(0.321)	(0.253)	(0.152)	(0.036)
3. Chemical products	370	2893	36.5	12.4	0.044	0.010	0.034	0.048	0.001	0.015	0.050	0.043	0.034
					(0.228)	(0.058)	(0.227)	(0.172)	(0.167)	(0.274)	(0.226)	(0.132)	(0.065)
4. Agric. and ind. machinery	194	1639	36.1	12.6	0.022	0.012	0.010	0.023	-0.012	-0.005	0.030	0.042	0.033
					(0.287)	(0.029)	(0.285)	(0.189)	(0.184)	(0.372)	(0.271)	(0.146)	(0.042)
5. Electrical goods	265	2032	31.0	18.5	0.037	0.007	0.029	0.038	-0.005	0.017	0.047	0.048	0.030
					(0.277)	(0.046)	(0.275)	(0.171)	(0.188)	(0.360)	(0.267)	(0.172)	(0.050)
6. Transport equipment	206	1692	39.1	13.7	0.029	0.007	0.021	0.036	-0.013	0.008	0.034	0.044	0.028
					(0.300)	(0.034)	(0.299)	(0.183)	(0.216)	(0.380)	(0.279)	(0.167)	(0.049)
7. Food, drink and tobacco	432	3404	33.1	10.4	0.037	0.020	0.017	0.044	0.000	0.005	0.044	0.046	0.037
					(0.218)	(0.056)	(0.220)	(0.182)	(0.168)	(0.290)	(0.230)	(0.165)	(0.063)
8. Textile, leather and shoes	395	2974	35.0	27.1	0.004	0.015	-0.011	0.023	-0.023	-0.024	0.013	0.048	0.031
					(0.231)	(0.038)	(0.230)	(0.189)	(0.177)	(0.336)	(0.229)	(0.172)	(0.043)
9. Timber and furniture	265	2017	54.6	22.4	0.002	0.017	-0.015	0.035	-0.015	-0.034	0.014	0.053	0.035
					(0.246)	(0.034)	(0.245)	(0.159)	(0.183)	(0.361)	(0.248)	(0.163)	(0.040)
10. Paper and printing products	238	1975	43.7	24.7	0.024	0.013	0.011	0.039	-0.009	-0.008	0.031	0.048	0.034
					(0.193)	(0.067)	(0.188)	(0.223)	(0.151)	(0.254)	(0.194)	(0.133)	(0.070)

Table A2 (continued)
Descriptive statistics

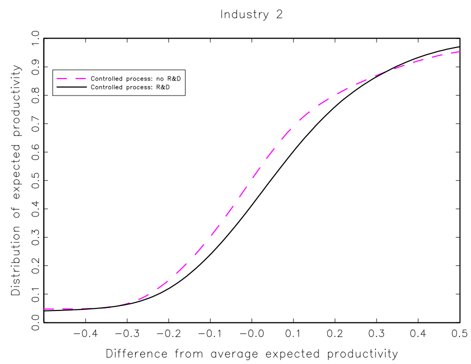
Industry	Advertising growth (s.d.)	With R&D % Obs.	With R&D R&D inten. (s. d.)	Market dynamism (s. d.)	Age (s. d.)
	(1)	(2)	(3)	(4)	(5)
1. Metals and metal products	-0.015	31.9	0.013	0.518	22.4
	(0.941)		(0.018)	(0.368)	(12.3)
2. Non-metallic minerals	-0.022	31.2	0.010	0.495	23.3
	(0.888)		(0.020)	(0.367)	(11.8)
3. Chemical products	0.004	55.0	0.025	0.536	26.6
	(0.806)		(0.034)	(0.351)	(12.8)
4. Agric. and ind. machinery	-0.016	53.7	0.027	0.519	24.9
	(0.821)		(0.031)	(0.364)	(12.3)
5. Electrical goods	0.003	58.9	0.032	0.522	22.6
	(0.843)		(0.043)	(0.371)	(12.0)
6. Transport equipment	-0.014	55.1	0.028	0.511	23.5
	(0.843)		(0.045)	(0.382)	(12.8)
7. Food, drink and tobacco	0.023	28.3	0.008	0.517	24.7
	(0.842)		(0.022)	(0.324)	(12.2)
8. Textile, leather and shoes	0.023	27.5	0.017	0.399	21.6
	(0.838)		(0.028)	(0.347)	(12.3)
9. Timber and furniture	-0.003	21.4	0.011	0.455	17.5
	(0.940)		(0.023)	(0.354)	(10.8)
10. Paper and printing products	-0.003	16.4	0.015	0.472	22.5
	(0.850)		(0.026)	(0.338)	(12.4)

Figure 1: Distribution of expected productivity: DJ model

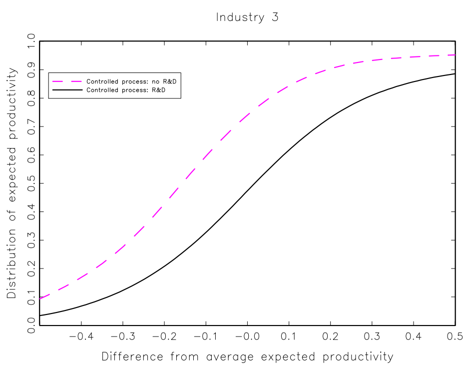
Metals and metal products



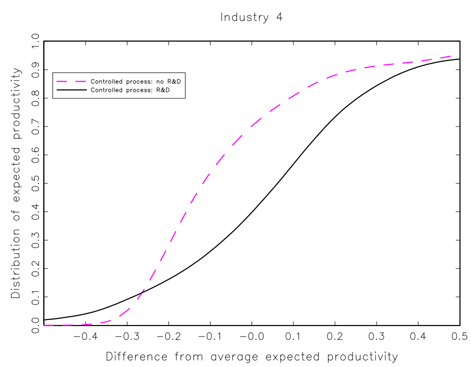
Non-metallic minerals



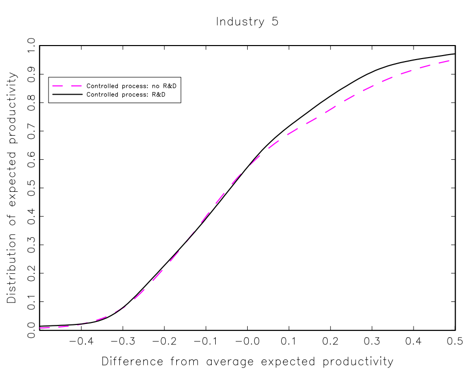
Chemical products



Agric. and ind. machinery



Electrical goods



Transport equipement

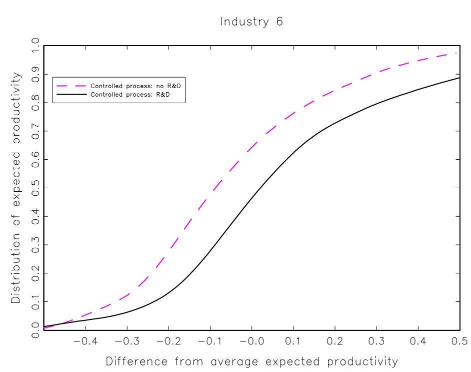
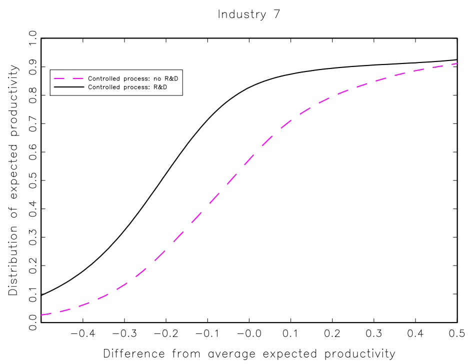
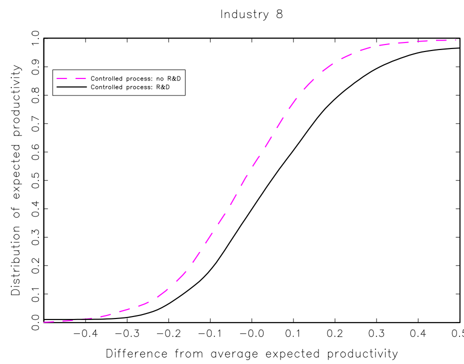


Figure 1: Distribution of expected productivity: DJ model (cont.)

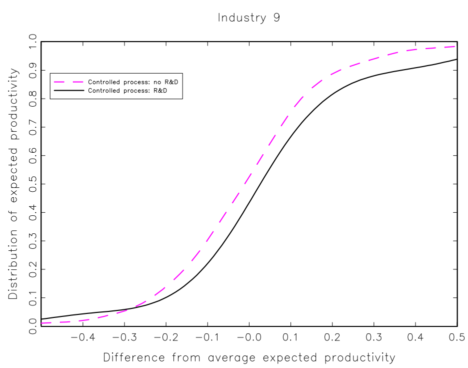
Food, drink and tobacco



Textile, leather and shoes



Timber and furniture



Paper and printing products

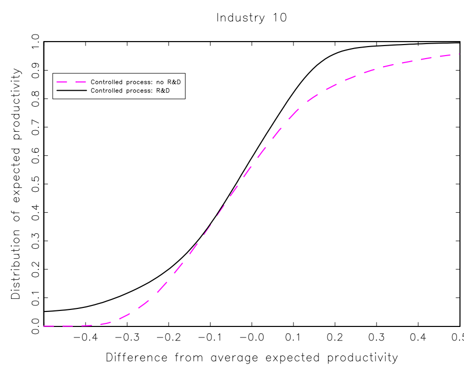
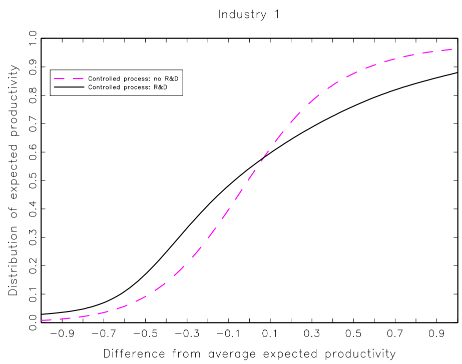
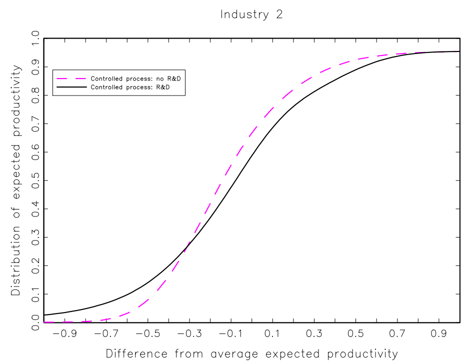


Figure 2: Distribution of expected productivity: no prices model

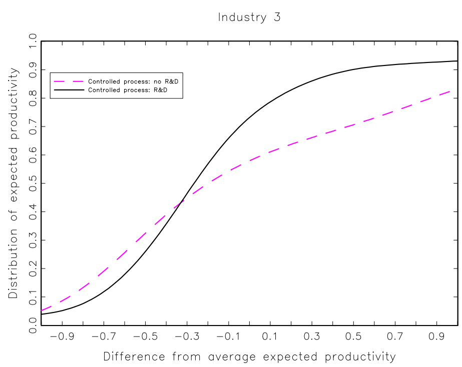
Metals and metal products



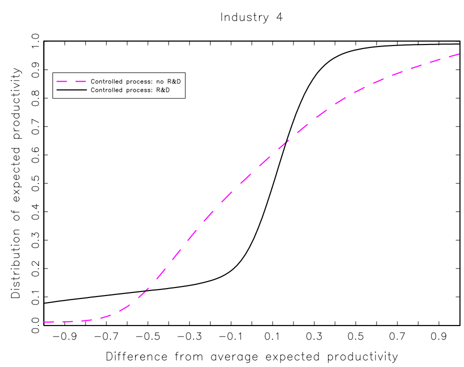
Non-metallic minerals



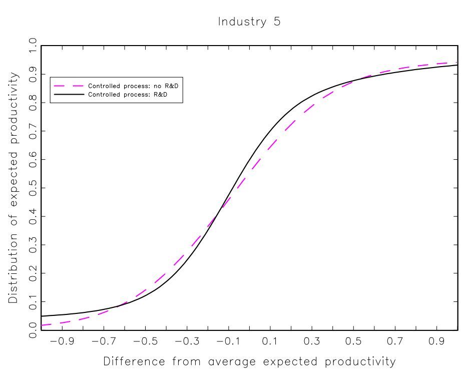
Chemical products



Agric. and ind. machinery



Electrical goods



Transport equipement

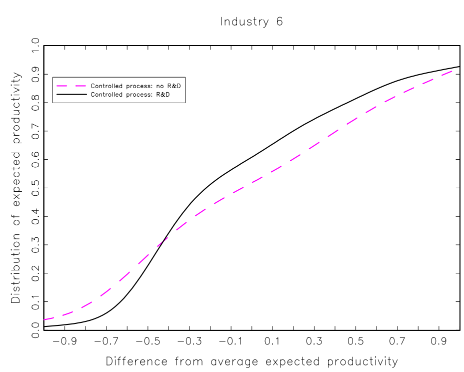
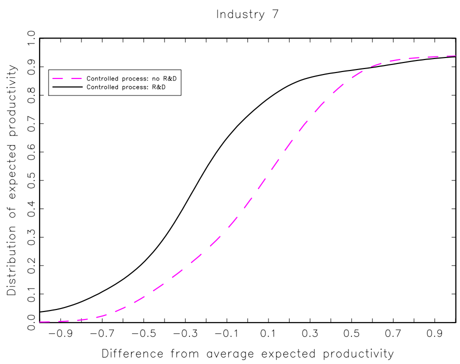
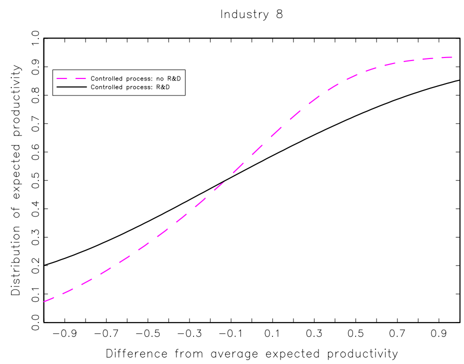


Figure 2: Distribution of expected productivity: no prices model (cont.)

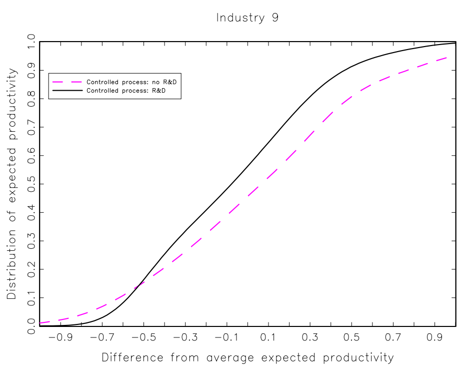
Food, drink and tobacco



Textile, leather and shoes



Timber and furniture



Paper and printing products