Reconsidering learning by exporting*

Miguel Manjón[a], Juan A. Máñez[b],

María E. Rochina-Barrachina[b] and Juan A. Sanchis-Llopis[b]

[a] QURE-CREIP Department of Economics, Universitat Rovira i Virgili

[b] Department of Applied Economics II and ERICES, Universitat de València

Abstract

Self-selection and learning-by-exporting are the main explanations for the higher productivity of exporting firms. But, whereas evidence on self-selection is largely undisputed, results on learning-by-exporting are mixed and far from conclusive. However, recent research (De Loecker, 2010) has shown that the conclusions from previous learning-by-exporting studies may have been driven by strong assumptions about the evolution of productivity and the role of export status. Relaxing these assumptions turns out to be critical to find evidence of learning-by-exporting in a representative sample of Spanish manufacturing firms. Our results indicate that the yearly average gains in productivity are around 3% for at least four years.

Keywords: learning-by-exporting, productivity

The authors acknowledge financial support from the Ministerio de Ciencia e Innovación (projects ECO2008-04576/ECON, ECO2008-06395-C05-03 and SEJ2010-19088/ECON), the Generalitat Valenciana (project PROMETEO/068) and the Generalitat de Catalunya ("Xarxa de Referència d'R+D+I en Economia i Polítiques Públiques" and the 2009-SGR-322 Program). Usual caveats apply.

1. Introduction.

The relation between exports and productivity has been extensively studied.[1] Using rich micro datasets from a wide range of different countries, this research has consistently found that exporters are generally more productive than non-exporters. Also, there is general agreement that the main mechanisms that generate these productivity gains from export market participation act ex-ante (self-selection) and/or ex-post (learning-by-doing, LBE hereafter).

In the self-selection mechanism firms need to reach a minimum productivity threshold to enter the more competitive foreign markets (Melitz 2003). Thus, only the *ex-ante* more productive firms are able to sell abroad. On the other hand, in the LBE mechanism firms improve their productivity after entering a foreign market (Clerides *et al.*, 1998). Therefore, exporting results in productivity gains because, among others, the growth in sales allows firms to profit from economies of scale, knowledge flows from international customers provide information about process and product innovations that might reduce costs and, improve quality, and increased competition forces firms to behave more efficiently.

Whereas there exists widespread empirical evidence supporting the hypothesis of self-selection into export markets (see e.g. Bernard and Jensen, 1999; ISGEP, 2008), the evidence on LBE is mixed and far from conclusive.[2] Some works do not find any evidence of post-entry productivity changes (Wagner, 2002,

[1] See Greenaway and Kneller (2007a) and Wagner (2007 and 2011) for thorough reviews of this literature.

[2] Silva *et al.* (2010) provide a detailed survey of the learning by exporting literature. Further, Martins and Yang (2009) provide a meta-analysis of 33 empirical studies. Singh (2010) concludes that studies supporting self-selection overwhelm studies supporting learning-by-exporting.

Arnold and Hussinger, 2005, Hansson and Lundin, 2004), and those that find differ in the time span and extent of the productivity changes (Greenaway and Kneller, 2004, 2007b, 2008; Girma *et al.*, 2004; Van Biesebroeck, 2005; De Loecker, 2007, 2010; Serti and Tomassi, 2008; Máñez *et al.*, 2010).

As De Loecker (2010) has recently shown, however, most previous tests on the existence of the LBE mechanism could be flawed. The usual empirical strategy is to look at whether a productivity estimate, typically obtained as the residual of a production function estimation, increases after firms enter in the export market. But for such an estimate to make sense, past export experience should be allowed to impact future productivity. Yet some previous studies (implicitly) assume that the productivity term in the production function specification is just an idiosyncratic shock (Wagner, 2002; Hansson and Lundin, 2004; Greenaway and Kneller, 2004, 2007b, 2008; Girma *et al.*, 2004; Máñez *et al.*, 2010), while others assume that this term is governed by an exogenous Markov process (Arnold and Hussinger, 2005; Serti and Tomassi, 2008). It is this sort of assumptions, often critical to obtain consistent estimates (Ackerberg *et al.*, 2007), what makes these tests of the existence of LBE to lack internal consistency.

In the first part of our analysis, we address this drawback by considering a more general process driving the law of motion for productivity. In particular, we explore the potential role that export experience might have in shaping firm's future productivity. Moreover, in the specification of the production function we acknowledge that exporters and non-exporters may have different demands of materials. Lastly, we incorporate these features into the GMM framework proposed by Wooldridge (2009).

In the second part of our empirical analysis we analyse whether allowing past export experience to affect productivity has any impact in the analysis of LBE. Thus, we use matching techniques to analyse the causal effects of exports on productivity, both using a productivity estimate based on an exogenous Markov process and a productivity estimate based on a more general process in which we allow past export experience to affect productivity. We do not obtain evidence of learning-by-exporting when using productivity estimates based on an exogenous Markov process. In contrast, we obtain yearly average gains in productivity of around 3% when estimates are based on the more general process. The international comparability of our results is difficult as to the best of our knowledge De Loecker (2010) is the only research that uses a similar method as in this paper.

Our approach is essentially the same as that followed by De Loecker (2007, 2010); see also Van Biesebroeck (2005). In particular, Van Biesebroeck (2005) is probably the first study to extend the estimation framework developed by Olley and Pakes (1996) to include lagged export participation status as a state variable in the estimation of productivity. However, De Loecker (2007) includes export status in the control function, and he is the first to explicitly consider two different demands of investment for exporters and non-exporters. De Loecker (2007, 2010) allows the law of motion of productivity to depend on past export status.[3] In our study we will consider, as in De Loecker (2010), both two different demands of

[3] See also Kasahara and Rodrigue, 2008, for an analogous approach to De Loecker (2010) but focusing on the analysis of the effect of imports on productivity.

intermediate materials and we will allow export experience in the law of motion of productivity.

We differ from De Loecker (2007, 2010) in that he uses investment as a proxy variable (and so does Van Biesebroeck, 2005) whereas we use intermediate materials. In this way we avoid possible concerns about zero-investment observations (Levinshon and Petrin, 2003) and the invertibility of the investment function (Van Biesebroeck, 2005). Further, whereas De Loecker (2007) uses an export participation dummy to proxy for export experience, we will use exports sales. Notice that by using export sales we control for both the decision to export and the intensity of exporting. Another important difference with respect to these related studies arises from the data. Whereas we analyse a panel of firms observed for almost twenty years, De Loecker (2007, 2010) observe firms for 7 years at most and Van Biesebroeck (2005) for less than three. Having a longer time period should help to identify LBE effects. It is also interesting to note that Van Biesebroeck (2005) and De Loecker (2007, 2010) provide evidence from developing countries (at the observational period), whereas we do from a developed country.

The rest of the paper is organized as follows. Section 2 briefly presents the data and provides empirical evidence of the significant differences between exporters and non-exporters in critical variables. Section 3 is devoted to the production function estimation method. Section 4 empirically analyses the relationship between firms' productivity and their export status. Section 5 concludes.

2. Data and descriptive analysis of exporters *versus* non-exporters.

The data used in this paper are drawn from the *Encuesta sobre Estrategias Empresariales* (ESEE, hereafter) for the period 1990-2008. This is an annual survey that is representative of Spanish manufacturing firms classified by industrial sectors and size categories. It provides exhaustive information at the firm level, and its panel nature allows following firms over time.

The sampling procedure of the ESEE is the following. Firms with less than 10 employees were excluded from the survey. Firms with 10 to 200 employees were randomly sampled, holding around 5% of the population in 1990. All firms with more than 200 employees were requested to participate, obtaining a participation rate around 70% in 1990. Important efforts have been made to minimise attrition and to annually incorporate new firms with the same sampling criteria as in the base year, so that the sample of firms remains representative over time.[4] Our final sample is an unbalanced panel of 2142 manufacturing firms (15,774 observations) that provided information for at least three consecutive years over the period 1990 to 2008 and without missing information on critical variables for the analysis, such as firm output, capital, intermediates and labour, among others.[5]

The firm export status is computed using the following question in the ESEE: "Indicate whether the firm, either directly, or through other firms from the

[4] See http://www.funep.es/esee/ing/i_esee.asp for further details.

[5] We do not use any observation for 1990 as we cannot compute productivity for this year in this survey.

same group, has exported during this year (including exports to the European Union) and its value".

As regards the firm export activity, Table 1 reports both the export participation and export intensity rates by industry in our sample for the period 1991-2008.[6] From these statistics we observe that the proportion of exporting firms for all industries is 58.24%, although there are significant differences across industries (the highest export participation rate, 84.65%, corresponds to the industry *Transport equipment*, and the lowest to the industry *Food, drink and tobacco*, 41.69%).[7] As regards export intensity (exports over sales), we observe an average of 27.45% for all industries, with also significant differences across industries (*Transport equipment* has the highest export intensity rate, 43.58%, and *Paper and printing products* has the lowest, 14.29%).

[Table 1 around here]

Next, we identify some stylized facts about exporting and non-exporting firms, using a simple regression analysis (see Table 2). The objective of this analysis is to explore the relationship between firm exporting status and some basic firm characteristics (Bernard and Jensen, 1999, De Loecker, 2007). In particular, we estimate equations of the form:

---

[6] We consider the original industry classification of the ESEE summarised in nine industries to guarantee enough observations *per* industry (we use the same nine industries classification than in Doraszelski and Jaumandreu, 2009). A higher disaggregation makes unfeasible industry-by-industry productivity estimation.

[7] The results of export participation and export intensity could be biased by the fact that the ESEE only surveys firms with more than 10 workers.

$$\log\left(y_{it}\right)=\beta_0+\beta_1 export_{it}+\delta\log\left(size_{it}\right)+\sum_{i=2}^{9}\gamma_i ind_i+\sum_{t=1994}^{2008}\lambda_t year_t+e_{it} \qquad (1)$$

where the dependent variable y_{it} is alternatively output, capital and intermediate materials per worker, size (as measured by the number of employees) and age. The variable $export_{it}$ is a dummy variable that takes value one if the firm exports and zero otherwise. We also control for size (except for the size regression), industries and years.

[Table 2 around here]

The differences (in %) between exporters and non-exporters for each of the five considered firm characteristics, computed from the estimated coefficient β_1 as $100(\exp(\beta_1) - 1)$, show that exporters are significantly bigger, older, more capital and intermediate materials intensive and have larger labour productivity than non-exporters. Consequently, it seems important to acknowledge the significant differences between exporters and non-exporters when estimating productivity. We do this by considering that exporters have a different demand function for intermediate materials than non-exporters. As pointed out by De Loecker (2007, 2010), this might be an important refinement in the analysis of the learning-by-exporting effect.

3. Production function estimation.

We assume that firms produce a homogenous good using a Cobb-Douglas technology:

$$y_{it} = \beta_0 + \beta_l l_{it} + \beta_k k_{it} + \beta_a a_{it} + \beta_m m_{it} + \omega_{it} + \eta_{it} \tag{2}$$

where y_{it} is the natural log of production of firm *i* at time *t*, l_{it} is the natural log of labour, k_{it} is the natural log of capital, a_{it} is the natural log of the age of the firm and, m_{it} is the natural log of intermediate materials. As for the unobservables, ω_{it} is the productivity (not observed by the econometrician but observable or predictable by firms) and η_{it} is a standard *i.i.d.* error term that is neither observed nor predictable by the firm.

It is also assumed that age and capital evolve following a certain law of motion that is not directly related to current productivity shocks (i.e. they are state variables), whereas labour and intermediate materials are inputs that can easily be adjusted whenever the firm faces a productivity shock (i.e. they are freely variable factors).[8]

Under these assumptions, Olley and Pakes (1996, OP hereafter) show how to obtain consistent estimates of the production function coefficients using a semiparametric procedure; see also Levinshon and Petrin, 2003, (LP hereafter) for a closely related estimation strategy. However, here we follow Wooldridge (2009),

[8] The law of motion for capital follows a deterministic dynamic process according to which $k_{it} = (1-\delta)k_{it-1} + I_{it-1}$. Thus, it is assumed that the capital the firm uses in period *t* was actually decided in period *t-1* (it takes a full production period for the capital to be ordered, received and installed by the firm before it becomes operative). The age of the firm is also considered as a deterministic state variable that evolves according to $a_{it} = a_{it-1} + 1$. Labour and materials (unlike capital) are chosen in period *t*, the period they actually get used (and, therefore, they can be a function of ω_{it}). These timing assumptions make them non-dynamic inputs, in the sense that (and again unlike capital) current choices for them have no impact on future choices.

who argues that both OP and LP estimation methods can be reconsidered as consisting of two equations which can be jointly estimated by GMM: the first equation tackles the problem of endogeneity of the non-dynamic inputs (that is, the freely variable factors); and, the second equation deals with the issue of the law of motion of productivity. Next we consider each in detail.

Let us start considering first the problem of endogeneity of the non-dynamic inputs. Correlation between labour and intermediate inputs with productivity complicates the estimation of equation (2), because it makes the OLS estimator biased and the fixed-effects and instrumental variables methods generally unreliable (Ackerberg *et al.*, 2007). Both OP and LP methods use a control function approach to solve this problem, by using investment in capital and materials, respectively, to proxy for "unobserved" firm productivity.

In particular, the OP method assumes that the demand for investment in capital, $i_{it} = i_t\left(k_{it}, a_{it}, \omega_{it}\right)$, is a function of firms' capital, age and productivity. To circumvent the problem of firms with zero investment in capital, the LP method uses the demand for materials, $m_{it} = m_t\left(k_{it}, a_{it}, \omega_{it}\right)$, instead, as a proxy variable to recover "unobserved" firm productivity. Since we follow this last approach, we concentrate on the demand of materials hereafter.[9]

[9] Both the investment of capital demand function and the demand for intermediate materials are assumed to be strictly increasing in ω_{it} (in the case of the investment of capital this is assumed in the region in which i_{it}>0). That is, conditional on k_{it} and a_{it}, a firm with higher ω_{it} optimally invests more (or demands more materials).

Therefore, when estimating productivity using these general versions of OP and LP, in a sample with exporters and non-exporters, it is assumed that the demand of intermediate materials both for exporters and non-exporters is identical. However, exporters differ in many characteristics from non-exporters (see Bernard and Jensen, 1999, for a classical reference; and Table 2 in this paper). Thus, analogously to De Loecker (2007, 2010), when analysing the LBE hypothesis we consider different demands of intermediate materials for exporters and non-exporters. Thus, we write the demand of materials as:

$$m_{it} = m_E\left(k_{it}, a_{it}, \omega_{it}\right) \tag{3}$$

where we include the subscript E to denote different demands of intermediate inputs for exporters and non-exporters. Also, since the demand of intermediate materials is assumed to be monotonic in productivity, it can be inverted to generate the following inverse demand function for materials:

$$\omega_{it} = h_E\left(k_{it}, a_{it}, m_{it}\right) \tag{4}$$

where h_E is an unknown function of k_{it}, a_{it} and m_{it}. Then, substituting the above expression (4) into the production function (2) we get:

$$y_{it} = \beta_0 + \beta_l l_{it} + \beta_k k_{it} + \beta_a a_{it} + \beta_m m_{it} + h_E\left(k_{it}, m_{it}, a_{it}\right) + \eta_{it} \tag{5}$$

Finally, by considering two different demand functions for intermediate materials (one for exporters and another one for non-exporters), our first estimation equation results in:

$$y_{it} = \beta_l l_{it} + 1(non-\exp)H_0\left(k_{it}, a_{it}, m_{it}\right) + 1(\exp)H_1\left(k_{it}, a_{it}, m_{it}, E_{it}\right) + \eta_{it} \tag{6}$$

where 1(*non-exp*) and 1(*exp*) are indicator functions that take value one for non-exporters and exporters, respectively. Notice that whereas the demand function of intermediate materials for non-exporters depends on k_{it}, m_{it} and a_{it}, this demand also depends on export experience E_{it} in the case of exporters.

As for the timing of the firm's export decision, following Van Biesebroeck (2005), we assume that the firm decides whether to export or not in period t knowing its productivity in t-1 (but not in period t). Therefore, the firm export experience in a given period only affects its productivity level in the next period. In particular, we proxy firms' export experience (E_{it}) in equation (6) by export sales in period t-1. Doing so solves the potential simultaneity that could arise if firms take their export decision in t, after observing their productivity (ω_{it}).

Further, the unknown functions H_0 and H_1 in (6) are going to be proxied by third degree polynomials in their respective arguments. Notice, however, that we cannot identify β_k, β_m and β_a from (6). This is achieved by the inclusion of a second estimation equation in the GMM-system that deals with the law of motion of productivity.

The standard OP/LP approaches consider that productivity evolves according to an exogenous Markov process:

$$\omega_{it} = E\left[\omega_{it} \middle| \omega_{it-1}\right] + \xi_{it} = f(\omega_{it-1}) + \xi_{it} \tag{7}$$

where f is an unknown function that relates productivity in t with productivity in t-1 and ξ_{it} is an innovation term uncorrelated by definition with k_{it} and a_{it}. However, this assumption neglects the possibility of previous export experience to affect productivity. Consequently, here we consider a more general (endogenous

Markov) process in which previous export experience can influence the dynamics of productivity:

$$\omega_{it} = E\left[\omega_{it} \middle| \omega_{it-1}, E_{it-1}\right] + \xi_{it} = f\left(\omega_{it-1}, E_{it-1}\right) + \xi_{it} \tag{8}$$

where $E_{\text{it-1}}$ could be a vector of variables summarising a firm's export experience, such as export participation, export sales, export intensity or the number of export markets the firm serves, among others (De Loecker, 2010). As already mentioned above, in our estimates we proxy export experience using export sales.

Let us now rewrite the production function (2) using (8) as:

$$y_{it} = \beta_0 + \beta_l l_{it} + \beta_k k_{it} + \beta_a a_{it} + \beta_m m_{it} + f\left(\omega_{it-1}, E_{it-1}\right) + \xi_{it} + \eta_{it} \tag{9}$$

Further, since $\omega_{it} = h_E\left(k_{it}, m_{it}, a_{it}\right)$, we can rewrite $f(\omega_{it-1}, E_{it-1})$ as:

$$\begin{aligned} f(\omega_{it-1}, E_{it-1}) &= f\left[h_E\left(k_{it-1}, a_{it-1}, m_{it-1}\right), E_{it-1}\right] = F_E\left(k_{it-1}, a_{it-1}, m_{it-1}\right) = \\ &= 1(non - \exp) F_0\left(k_{it-1}, a_{it-1}, m_{it-1}\right) + 1(\exp) F_1\left(k_{it-1}, a_{it-1}, m_{it-1}, E_{it-1}\right) \end{aligned} \tag{10}$$

with F_0 and F_1 being unknown functions to be proxied by third degree polynomials in their respective arguments.

Lastly, substituting (10) into (9), our second estimation equation is given by:

$$\begin{aligned} y_{it} &= \beta_l l_{it} + \beta_k k_{it} + \beta_a a_{it} + \beta_m m_{it} + 1(non - \exp) F_0\left(k_{it-1}, a_{it-1}, m_{it-1}\right) + \\ &+ 1(\exp) F_0\left(k_{it-1}, a_{it-1}, m_{it-1}, E_{it-1}\right) + u_{it} \end{aligned} \tag{11}$$

where $u_{it}=\xi_{it}+\eta_{it}$ is a composed error term.

Wooldridge (2009) proposes to estimate jointly equations (6) and (11) by GMM using the appropriate instruments and moment conditions for each equation.

This joint estimation strategy has several advantages: i) it increases efficiency relatively to the two step traditional procedures; ii) it makes unnecessary to do bootstrapping for the calculus of standard errors; and, iii) it solves the problem, pointed out by Ackerberg *et al.* (2006), of identification of the labour coefficient in the separate estimation of equation (6). This procedure allows us to obtain, *per* each one of the nine industries considered, both coefficient estimates of the production function and firms' productivity estimates. In particular, to estimate firms' productivity, both assuming and exogenous and an endogenous Markov process, we use the corresponding polynomial approximation of expression (10).[10]

4. The relationship between the export status and productivity.

Having estimated firms' productivity, both assuming an exogenous and an endogenous Markov process, next we use these estimates to test for learning-by-exporting. Prior to this, however, we check for each industry whether exporters are more productive than non-exporters using stochastic dominance techniques. This will provide a first picture of the effects of exporting on productivity.

For this comparison we define as exporters those firms that export at least one year along the years they are in the sample and as non-exporters those firms that do not export in any of the sample years. Also, since the Kolmogorov-Smirnov one and two-sided tests of stochastic dominance (KS, hereafter) require independence of observations both between the two samples under comparison and among the observations of a given sample, analogously to Doraszelski and

[10] These estimates are available upon request.

Jaumandreu (2009) for R&D, for each industry *j* we compare:

$$F_j\,(productivity) = G_j\,(productivity), \quad j = 1,...,9 \tag{12}$$

where F is the cumulative distribution function of the average productivity for exporters (calculated as the average over the years they export) and G is the cumulative distribution function of the average productivity for non-exporters (calculated as the average over the years they are in the sample).[11]

[Tables 3 and 4 around here]

Tables 3 and 4 report the KS tests when considering an exogenous Markov process and an endogenous Markov process, respectively. Results show that, irrespectively of the type of Markov process considered, we reject the null hypothesis of equality of the two distributions (at a 5% or even much smaller significance level) in all industries. Furthermore, we can never reject the null that the productivity of exporters is higher than that of non-exporters. Therefore, the productivity distribution for exporters stochastically dominates that of non-exporters.

The analysis of LBE implies to test whether export participation has any impact on export productivity growth. However, comparing the productivity growth of export starters and non-exporters after the former start to export does not allow assessing if the observed differences are due to LBE or to self-selection. But to properly control for the direction of causality from exporting to productivity growth, we would need to compare the actual productivity growth of

[11] See Delgado *et al.* (2002) for a more detail description of the application of the KS tests of stochastic dominance and Máñez *et al.* (2010) for a panel data application (analogous to ours).

export starters after starting to export with the productivity growth of the same firms if they would not have started to export. The problem is that we do not have information about the counterfactual situation: the productivity growth of export starters if they would not have started to export. Matching techniques provide a way to construct this counterfactual.

To classify a firm as an export starter at period t we require two conditions: (i) the firm should not have exported during the sample periods previous to t; and, (ii) it should be observed in the sample at least for two years previous to t. We classify as non-exporters those firms that do not export during the whole period they are in the sample and for which we have information in the sample for at least two years.

More formally, let Δy_{it}^{D} denote the growth rate of productivity and $D_{it} \in \{0, 1\}$ be an indicator of whether firm i is an export starter (i.e., a firm that exports for the first time in the sample years) at period t (as opposed to a non-exporter). Thus, let Δy_{it+s}^{1} be the productivity growth between t and $t+s$ for an export starter (with s>0 and firm i being classified as an export starter in t) and let Δy_{it+s}^{0} be the growth outcome for firm i had it not started to export in t. Using this notation, the causal effect of exporting for firm i at $t+s$ can be defined as

$$\Delta y_{it+s}^{1} - \Delta y_{it+s}^{0} \tag{13}$$

Also, following the policy/treatment evaluation literature (see Heckman *et al.*, 1997), we can define the average effect of exporting on firms who start to export as

$$E\left(\Delta y^1_{it+s} - \Delta y^0_{it+s} \middle| D_{it} = 1\right) = E\left(\Delta y^1_{it+s} \middle| D_{it} = 1\right) - E\left(\Delta y^0_{it+s} \middle| D_{it} = 1\right) \tag{14}$$

The main problem of using (14) for making causal inference is that the counterfactual Δy^0_{it+s} is not observed (notice that this is the average productivity growth that export starters would have experienced had they not started to export). We overcome this problem using matching techniques to identify, among the pool of non-exporters in *t,* those with a distribution of observable variables (*X* in *t-1*) affecting productivity growth and the probability of exporting, as similar as possible to that of export starters. It is then assumed that, conditional on *X*, firms with the same characteristics are randomly exposed to export activities. Thus, expression (14) can be rewritten as

$$E\left(\Delta y^1_{it+s} - \Delta y^0_{it+s} \middle| D_{it} = 1\right) = E\left(\Delta y^1_{it+s} \middle| X_{it-1}, D_{it} = 1\right) - E\left(\Delta y^0_{it+s} \middle| X_{it-1}, D_{it} = 0\right) \tag{15}$$

Since the set of observable variables that may potentially affect the firms' probability of exporting and their productivity growth is quite large, the question that arises is what is the appropriate variable to match firms (and in case of using more than one variable, what are the appropriate weights). We deal with this issue using the propensity score techniques proposed by Rosenbaum and Robin (1983). This means that if the decision to start exporting is a random process conditioning upon *X*, it is also random conditioning on the probability of exporting.

Therefore, before performing the matching we obtain the probability of becoming an export starter (i.e., the propensity score) as the predicted probability of the following probit model

$$P(D_{it} = 1) = \Phi\{\omega_{it-1}, k_{it-1}, a_{it-1}, s_{it-1}, ind, year\} \tag{16}$$

where $\Phi(\cdot)$ is the normal cumulative distribution function and the set of observable characteristics included in the model are lagged productivity, capital, age, size, and industry and year dummies.

We then use nearest neighbours matching to construct the counterfactual (Becker and Ichino, 2002). In particular, matching is performed using the *psmatch2* command of Stata (Leuven and Sianesi, 2003). Notice that imposing the balancing constraints ensures that the mean propensity score is not different between export starters and the matched non-exporters. We will therefore match firms on the basis of the probability to export for the first time (as captured by the regressors).[12]

[Table 5 around here]

We report results from the matching analysis in Table 5. More specifically, we provide estimates of the extra-productivity growth (EPG) for export starters both under the assumption of a pure exogenous productivity process and under a law of motion for productivity in which we allow for past export experience to impact future productivity (endogenous Markov process). To analyse the time

[12] Abadie and Imbens (2008) show that due to the extreme non-smoothness of nearest neighbours matching, the standard conditions for bootstrapped standard errors are not satisfied, leading the bootstrap variance to diverge from the actual variance. This may be corrected either by subsampling (Politis *et al.*, 1999) or using the Stata *nnmatch* command (Abadie *et al.*, 2004). We report results from both approaches (see Table 5) and do not find substantial differences between the estimated standard errors. Also, we evaluate the quality of the matching analysis using alternative indicators of the resulting balancing of the observable variables within the matched samples (see Appendix for details). Our analysis indicates that the matching procedures performed well.

span of EPG of export starters, we provide EPG estimates for the periods t and $t+s$ for $s = 1,..., 4$.

Our results indicate that allowing for past-export experience to affect productivity crucially determines the analysis of the LBE hypothesis. This is so as we do not get any significant EPG for any of the periods considered when we assume an exogenous Markov process for productivity. However, we obtain highly significant and positive EPG for all periods when we consider that past export experience affects productivity, i.e., when we allow for an endogenous Markov process. In particular, we obtain an extra yearly average cumulative growth rate of around 3% for the productivity of an exporter after entry in relation to a non-exporter. Furthermore, this average extra productivity growth is obtained since the first year of exporting.

Our estimates of the effect of starting to export on cumulative productivity growth are somehow modest if we compare them to the estimates obtained by De Loecker (2007), the only paper that uses a similar methodology. In particular, for periods $s = 1$, 2, 3 and 4 our estimates are 3.6, 5.4, 10.8 and 14.4 %, respectively, whereas those of De Loecker (2007), when he allows both different demand for investments for exporters and non-exporters and introduces export status in the law of motion of productivity, for $s = 1,2,3$ and 4 are 14.7, 27.3, 41.4 and 30.6%, respectively.[13] This difference in the results could be due to the fact that, by the time of the analysis, Slovenia was considered a developing economy and, therefore, the scope of learning from participation in international markets was higher as

[13] These estimates are reported in Appendix B in page 97 in De Loecker (2007).

compared to the case of an economy fully integrated in international markets like Spain.

5. Conclusions.

The extensive literature analysing the relationship between exports and productivity has concluded that exporters are generally more productive than non-exporters and that only the ex-ante more efficient firms enter into export markets (i.e., there is self-selection into export markets). However, the higher productivity of exporters could be also the result of learning-by-exporting. Recent research has shown that previous empirical studies in this area have imposed strong assumptions about the evolution of productivity and the role of export status that may have biased the estimates towards the rejection of the learning-by-exporting hypothesis.

We investigate this tenet using a two-step strategy. In the first step we use a Cobb-Douglas production function to estimate firm productivity by GMM. In particular, in the specification of the production function we consider that exporters and non-exporters have different demands of intermediate materials. We also assume that firms' expectations about their future productivity depend not only on their current productivity but also on their export experience. In a second step, we test for LBE using matching analysis. We find that relaxing the assumptions about the evolution of productivity and the role of export status are critical to obtain evidence of learning-by-exporting.

In particular, results from the matching procedure indicate that the productivity gains of exporting by Spanish manufacturing firms are not negligible.

However, our estimates are lower than those find by De Loecker (2007) for Slovenia.

References.

Abadie, A., Drukker, D., Herr, J and G. W. Imbens (2004), Implementing matching estimators for average treatment effects in Stata. The Stata Journal, 4(3), 290-311.

Abadie, A. and G.W. Imbens (2008), On the Failure of the Bootstrap for Matching Estimators. Econometrica, 76, 6, 1537-1557.

Ackerberg, D. A., K. Caves and G. Frazer (2006), Structural identification of production functions, Working Paper, Department of Economics, UCLA.

Ackerberg, D., L. Benkard, S. Berry and A. Pakes (2007), Econometric tools for analyzing market outcomes, in J. Heckman and E. Leamer (Eds.), Handbook of Econometrics, Vol. 6A. Amsterdam: North Holland.

Arnold, J. and K. Hussinger (2005), Export Behavior and Firm Productivity in German Manufacturing: A Firm-level Analysis, Review of World Economics / Weltwirtschaftliches Archiv, 141, 2, 219–43.

Becker, S., and A. Ichino (2002), Estimation of average treatment effects based on propensity scores. Stata Journal, 2, 358-377.

Bernard, A. B. and J. B. Jensen (1999), Exceptional Exporter Performance: Cause, Effect, or Both? Journal of International Economics, 47(1), 1–25.

Clerides, S. K., S. Lach and J.R. Tybout (1998), Is Learning by Exporting Important? Micro-Dynamic Evidence from Colombia, Mexico, and Morocco, Quarterly journal of Economics, 113(3), 903–947.

De Loecker, J. (2007), Do Exports Generate Higher Productivity? Evidence from Slovenia. Journal of International Economics, 73, 1, 69–98.

De Loecker, J. (2010), A Note on Detecting Learning by Exporting," NBER Working Papers 16548, National Bureau of Economic Research, Inc.

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

Doraszelski, U. and J. Jaumandreu (2009), R&D and Productivity: Estimating Endogenous Productivity, mimeo, Harvard University.

Girma, S., D. Greenaway and R. Kneller (2004), Does Exporting Increase Productivity? A Microeconometric Analysis of Matched Firms. Review of International Economics, 12, 5, 855–66.

Greenaway, D. and R. Kneller (2004), Exporting and Productivity in the UK. Oxford Review of Economic Policy, 20, 3, 358–71.

Greenaway, D. and R. Kneller (2007a), Firm Heterogeneity, Exporting and Foreign Direct Investment. Economic Journal, 117, 517, 134–61.

Greenaway, D. and R. Kneller (2007b), Industry Differences in the Effect of Export Market Entry: Learning by Exporting? Review of World Economics / Weltwirtschaftliches Archiv, 143, 3, 416–32.

Greenaway, D. and R. Kneller (2008), Exporting, Productivity and Agglomeration. European Economic Review, 52, 5, 919–39.

Hansson, P. and N. Lundin (2004), Exports as Indicator on or a Promoter of Successful Swedish Manufacturing Firms in the 1990s. Review of World Economics

/ Weltwirtschaftliches Archiv, 140, 3, 415–45.

International Study Group on Exports and Productivity (ISGEP) (2008). Understanding Cross-Country Differences in Exporter Premia: Comparable Evidence for 14 Countries. Review of World Economics (Weltwirtschaftliches Archiv), 144, 596-635.

Kasahara, H. and J. Rodrigue (2008), Does the use of imported intermediates increase productivity? Plant-level evidence, Journal of Development Economics 87, 106–118.

Leuven, E, and B. Sianesi (2003), PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing. Statistical Software Components S432001, Boston College Department of Economics, revised 28 Dec 2006.

Levinsohn, J. and A. Petrin (2003), Estimating production functions using inputs to control for unobservables. Review of Economic Studies 70, 317–342.

Mánez-Castillejo, J.A., M.E. Rochina-Barrachina and J.A. Sanchis-Llopis (2010), Does firm size affect self-selction and learning-by-exporting? The World Economy, 33 (3), 315-346.

Martins, P.S. and Y. Yang (2009), The impact of exporting on firm productivity: a meta-analysis of the learning-by-exporting hypothesis. Review of World Economics 145 (3), 431-445.

Melitz, M. (2003), The impact of trade on intra-industry reallocations and aggregate industry productivity. Econometrica 71 (6), 1695–1725.

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

Politis, D.N., J.P. Romano and M. Wolf (1999), Subsampling, Springer-Verlag, New York.

Rosenbaum, P.R. and D.B. Rubin (1985) Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. The American Statistician, 39 (1), 33 - 38.

Serti, F., and C. Tomasi (2008). Self-Selection and Post-Entry Effects of Exports: Evidence from Italian Manufacturing Firms. Review of World Economics/Weltwirtschaftliches Archiv, 144 (4), 660–694.

Sianesi, B. (2004), An evaluation of the Swedish system of active labor market programs in the 1990s. The Review of Economics and Statistics, 86(1), 133-155.

Silva, A., A.P. Africano and Ó. Afonso (2010), Learning-by- exporting: What we know and what we would like to know. Universidade de Porto FEP Working Papers N. 364, March.

Singh, T. (2010), Does International Trade Cause Economic Growth? A Survey. The World Economy, 33, 1517-1564.

Van Biesebroeck, J. (2005), Exporting Raises Productivity in Sub-Saharan Manufacturing Plants. Journal of International Economics, 67, 2, 373–91.

Wagner, J. (2002), The Causal Effects of Export on Firm Size and Labour Productivity: First Evidence from a Matching Approach, Economics Letters, 77(2), 287–92.

Wagner, J. (2007), Exports and Productivity: A Survey of the Evidence from Firm Level Data, The World Economy, 30(12), 60–82.

Wagner, J. (2011), International Trade and Firm Performance: A Survey of Empirical Studies since 2006, IZA DP No. 5916.

Wooldridge, J.M. (2009), On estimating firm-level production functions using proxy variables to control for unobservables, Economics Letters, 104, 112–114.

Table 1. Export participation and export intensity by industry.

Industry	Export participation	Export intensity
Metals and metal products	50.00	33.23
Non-metallic minerals	55.56	35.50
Chemical products	70.69	24.38
Agric. and ind. machinery	69.46	36.52
Transport equipment	84.65	43.58
Food, drink and Tobacco	41.69	17.04
Textile leather and shoes	51.51	26.24
Timber and furniture	53.43	16.29
Paper and printing products	47.27	14.29
All industries	58.24	27.45

Table 2. Differences between exporters and non-exporters.

	Differences in %	*p*-value
Output per worker	51.78	0.000
Capital per worker	46.77	0.000
Intermediate materials per worker	83.41	0.000
Size	324.77	0.000
Age	65.75	0.000

Table 3. Productivity levels: exogenous Markov process.

	Number		Distributions are equal		Distributions of exporters dominate	
Industry	**Exporters**	**Non-Exporters**	S_1	***p*-value**	S_2	***p*-value**
Metals and metals products	150	153	3.733	0.000	0.117	0.973
Non-metallic minerals	88	61	1.602	0.007	0.083	0.986
Chemical products	168	44	3.221	0.000	0.000	1.000
Agric. and ind. machinery	139	40	3.216	0.000	0.080	0.987
Transport equipment	108	19	2.529	0.000	0.000	1.000
Food, drink and tobacco	145	134	5.617	0.000	0.000	1.000
Textile, leather and shoes	150	141	5.283	0.000	0.000	1.000
Timber and furniture	111	90	3.186	0.000	0.049	0.995
Paper and printing products	102	79	2.776	0.000	0.000	1.000

Table 4. Productivity levels: endogenous Markov process.

	Number		Distributions are equal		Distributions of exporters dominate	
Industry	**Exporters**	**Non-Exporters**	S_1	***p*-value**	S_2	***p*-value**
Metals and metals products	150	153	3.851	0.000	0.118	0.972
Non-metallic minerals	88	61	1.459	0.018	0.808	0.271
Chemical products	168	44	1.767	0.002	0.147	0.958
Agric. and ind. machinery	139	40	2.891	0.000	0.000	1.000
Transport equipment	108	19	2.417	0.000	0.261	0.873
Food, drink and tobacco	145	134	5.790	0.000	0.000	1.000
Textile, leather and shoes	150	141	4.885	0.000	0.000	1.000
Timber and furniture	111	90	3.709	0.000	0.000	1.000
Paper and printing products	102	79	2.334	0.000	0.065	0.991

Table 5. Estimates of extra-productivity growth for export starters.

Period	Nearest Neighbours	Observations	Exogenous Markov process EPG	Exogenous Markov process s.e.	Endogenous Markov process EPG	Endogenous Markov process s.e.
t/t+1	SS	92(2904)	0.048	0.039	0.036**	0.016
	A&I		0.048	0.034	0.036***	0.014
t/t+2	SS	67(2830)	0.047	0.048	0.027**	0.012
	A&I		0.047	0.043	0.027***	0.011
t/t+3	SS	51(2781)	0.056	0.055	0.036***	0.012
	A&I		0.056	0.048	0.036***	0.010
t/t+4	SS	43(2761)	0.025	0.056	0.036***	0.012
	A&I		0.025	0.050	0.036***	0.010

Notes:

1. EPG stands for extra productivity growth of export starters over matched non-exporters.
2. A&I means that standard errors have been calculated using Abadie and Imbens (2008) correction.
3. SS means that, following Politis *et al.* (1999), standard errors have been calculated by sub-sampling (2000 draws).
4. In the observations column we report the number of export starters and the number of control observations in parentheses, imposing common support.
5. s.e. stands for standard errors.
6. *, **, *** indicates significance at 10%, 5% and 1% level, respectively.
7. We define an export starter in *t* is a firm that, previous to year *t*, has declared that has not exported for a minimum of two years. We define as non-exporters those firms that do not declare to export any year during the whole sample period. Note also that when an export starter stops exporting its TFP growth from *t* to *t+s* (with $0<s\leq 4$) is only computed up to the previous period in which the firm stops exporting.

Appendix.

There are several approaches to evaluate whether the matching procedure is able to balance the distribution of the relevant variables, both for exporters and matched non-exporters, when one conditions on the propensity score. Following Sianesi (2004), we report a pseudo R^2 test and a joint significance test as matching quality indicators in Tables A.1 and A.2, respectively. Moreover, following Rosenbaum and Rubin (1985) we report the median absolute standardised biases before and after matching in Table A.3.

In particular, Sianesi (2004) suggests re-estimating the propensity score on the matched sample (that is, only on exporters and matched non-exporters) and comparing the probit pseudo R^2 before and after the matching. Since the probit pseudo R^2 indicates how well the regressors X explain the probability of exporting, after matching there should be no systematic differences in the distribution of the regressors between both groups. Consequently, the pseudo R^2 should be fairly low when performed in the matched sample. As reported in Table A.1, we obtain very small values for the pseudo R^2 after matching for all the periods. In addition, Sianesi (2004) proposes a joint significance test of all the probit regressors before and after matching. The interpretation of this test is that the joint significance of the regressors should be rejected after matching but not before, and this is indeed the result we obtain (see Table A.2).

Another indicator used to assess the distance in marginal distributions of the X variables is the median bias, where the bias refers to median absolute standardised bias before and after matching. The median is calculated over all regressors. Following Rosenbaum and Rubin (1985), for a given regressor the standardised difference before matching is the difference of the sample means between exporters and non-exporters as a percentage of the square root of the average of the sample variances from the two sub samples (exporters and non-exporters, respectively). The standardised difference after matching is analogously calculated using the corresponding values for the matched samples. In our results we obtain a substantial reduction in the standardised bias that seems to be consistent with the results obtained in other empirical studies (see Table A.3). Still, one potential problem in interpreting the standardised bias approach is that there is no clear indicator of the success of the matching procedure.

Table A.1. Probit pseudo R^2.

	Exogenous Markov Process		Endogenous Markov Process	
	Before	**After**	**Before**	**After**
t/t+1	0.118	0.045	0.116	0.032
t/t+2	0.136	0.054	0.136	0.039
t/t+3	0.150	0.062	0.149	0.074
t/t+4	0.159	0.092	0.157	0.073

Notes: Probit pseudo R^2 for export starters on covariates before matching and in matched samples (after matching).

Table A.2. $P>\chi^2$ (LR test of joint significance of coefficients in the Probit regression).

	Exogenous Markov Process		Endogenous Markov Process	
	Before	**After**	**Before**	**After**
t/t+1	0.000	0.999	0.000	1.000
t/t+2	0.000	0.999	0.000	1.000
t/t+3	0.000	1.000	0.000	0.999
t/t+4	0.000	0.998	0.000	0.999

Notes: $P > \chi^2$ is the p-value of the likelihood-ratio test after matching. This a test of the hypothesis that the regressors are jointly insignificant, i.e., that they are well balanced in the two samples.

Table A.3. Median bias in the Probit regression.

	Exogenous Markov Process		Endogenous Markov Process	
	Before	**After**	**Before**	**After**
t/t+1	11.644	5.792	10.347	5.759
t/t+2	11.603	4.164	11.040	6.527
t/t+3	16.021	10.854	16.056	8.760
t/t+4	18.839	12.216	17.278	11.627

Notes: Median bias refers to median absolute standardised bias before and after matching.