Professor Moisă ALTĂR, PhD
E-mail: altarm@gmail.com
The Bucharest University of Economic Studies
The Romanian - American University– FINSYS
Ana-Maria CAZACU (BANCU), PhD Candidate
E-mail: ana02mary@yahoo.com
The Bucharest University of Economics Studies

TESTING SELF-SELECTION AND LEARNING BY EXPORTING HYPOTHESES. THE CASE OF ROMANIA

Abstract. *This paper provides empirical evidence of export-productivity link, using a comprehensive set of data on Romanian firms activating in ten business sectors. By using the semi-parametric estimation technique developed by Levinsohn and Petrin (2003) to estimate Total Factor Productivity (TFP)we find that exporters display a productivity advantage compared to their domestic counterparts. We seek to distinguish between self-selection and learning by exporting through matching techniques and we find that the most productive firms self-select in activities in international markets, including importing. Learning effects (when TFP is used as outcome) are present on a narrow sample of companies, although in most of the cases are positive and significant when using labour productivity. Stronger evidence for positive effects of international trade is found for importers, which suggests a possible learning by importing phenomenon manifested in Romanian.*

Keywords*: endogeneity, selection bias, TFP, semi-parametric models, matching techniques, propensity score*

JEL Classification: C23, D24, F10, F23, F61

1. Introduction and state of art

Both academics and policy makers has increasingly focused in recent years on understanding the role of firm-level competitiveness and efficiency indicators in determining aggregate results (Competitiveness Research Network, 2014). Particularly, assessing the main drivers of international sales and the impact of the increased international trade on micro-level performance is of vital importance for a small and open economy. The literature centred on the link between involvement in export activities and firms' performance is focused on testing two main hypotheses: self-selection and learning-by-exporting. Self-selection hypothesis relies on the idea that companies entering export markets have higher productivity prior the entry, while learning-by-exporting hypothesis assumes that after acceding to foreign markets, firms have access to better technologies and information from foreign costumers and contracts and improve their overall performance. Thus, creating conditions favourable to trade, especially exports, can be one of the most important ways to obtain knowledge from abroad.

The scope of the present paper is to test the direction of causality between non-financial companies' performance in terms of productivity and external trade activities in case of Romania. For illustration, the period 2004-2011 was analysed, as it captures the most important episodes of external trade adjustment. In the process, we apply alternative algorithms for estimating firm-level Total Factor Productivity (TFP) and we use matching techniques in order to test self-selection of most productive firms in exporting activities and potential learning effects that might result due to involvement in activities on international markets.

A large number of studies confirm the self-selection hypothesis, Lopez (2004) arguing that this might be a conscious process by which firms increase their productivity with the explicit purpose of becoming exporters and producing high-quality goods for external markets. Bernard and Wagner (1997) find evidence of firms' self-selection in export sales in German manufacturing sector. Exporters perform better than their non-exporters counterparts in the industry, their productivity advantage varying around 15-20%. In case of Poland, Hagemejer (2007) finds that a 10% increase of TFP causes an increase in the probability of exporting by 4%.Altomonte and Bekes (2009) using data for Hungary, investigate the relation between trading activities (importing, exporting or both) and productivity. The main conclusion is related to the presence of an important self-selection effect of the most productive firms, induced by the heterogeneous sunk costs of trade, for both importers and exporters.

On the other hand, evidence in favour of learning effects is less clear-cut. Little empirical results confirm the learning by exporting hypothesis in the case of Germany (Bernard and Wagner, 1997) and Belgium (Pisu, 2008). De Loecker (2007), on a panel of Slovenian manufacturing firms during 1994–2000, uses matched sampling techniques and reveals that the export entrants become, on average, around 9% more productive after they start exporting and that the productivity gap between exporters and non-exporters increases over time. Learning effects are manifested in 13 out of 16 industries, with heterogeneous magnitude and timing across sectors.

The structure of paper is the following: Section 2 presents the data and Section 3 the methodological aspects in estimating micro-level TFP, one of the main indicators used in for identifying the direction of causality between exporting and efficiency gains. Section 4 focuses on the results of the matching techniques, while the finding' summary and concluding remarks are displayed in the last part.

2. Data description

In order to test the relation between productivity indicators (including TFP) and exporting activities, we use a large panel of Romanian companies activating in 10business sectors (Mining, Food, beverage and tobacco; Textile, leather; Chemicals, pharmaceutical products; Basic metal/Metallurgic industry; Electrical and electronic products; Auto industry;Electricity; Construction; Telecommunications and other services).

Table 2. Sample used in the analysis

Sector	Total number of companies, out of which:	Exporters	Importers
1.Mining	1.312	<10	22
2.Food, beverage and tobacco	10.435	98	314
3.Textile, leather	9.475	1.136	349
4.Chemicals, pharmaceutical products	1.179	49	110
5.Basic metal/Metallurgic industry	7.265	387	234
6.Electrical and electronic products	1.678	203	163
7.Auto industry	2.570	321	201
8.Electricity	779	16	33
9.Construction	50.848	33	240
10.Telecommunications and other services	2.814	<10	37
Total	**88.355**	**2.252**	**1.703**

The firm-level data is provided by the Ministry of Public Finance, the National Institute of Statistics and The National Trade Register Office for 2004-2011, recorded annually. These databases comprise a wide variety of micro-level information. Firms' value added is computed as a difference between net sales adjusted for changes in inventories and material costs. Labour is represented by the number of employees, while capital is proxied as firm's tangible fixed assets. Material consumption, used as a proxy to control for productivity shocks, is measured as the cost of intermediate inputs. All variables are expressed in real terms, using deflators from Eurostat. Business sectors are derived from 2 digit industry in which firms operate (based on NACE rev.2).Ownership dummies indicate if the company has national, mixed or foreign owners1.Based on exports and imports' volume, dummy variables for each company indicating whether it is an exporter/importer are used2 (international trade data is available starting 2007).

[1]over 50% of total equity is domestic-owned/ equally shared by national and domestic shareholders/ over 50% of total equity is foreign-owned.

[2]Only companies reporting net exports or imports more than EUR 100,000 in each quarter during 2007- 2011 were taken into consideration in the analysis.

Only active companies were taken into account (with positive turnover). After cleaning the data, the unbalanced panel dataset contains more than 88,000 firms (Table 2). Most of the estimates were performed in Stata 11, but Matlab 7 and EViews7 were also used, especially in processing the data.

3. Total factor productivity

3.1. Methodological problems in estimating TFP

The first step in the analysis is to appropriately estimate the unobservable TFP, as accurate measurement is essential in productivity comparisons. In estimating TFP, whose roots can be found in a seminal paper by Solow (1957), at least two[3] methodological issues emerge: simultaneity and selection problems. The former is related to the relation existing between productivity and input demands, input choices being determined (in part) by firms' believes about their productivity that can be reached when those inputs will be used. Since it violates OLS procedure's requirements that the inputs are exogenous, the traditional OLS estimators will be biased (as they are affected by the simultaneity problem). The selection problem is linked to the fact that, traditionally, the estimation of production function parameters is made by using a balanced panel. If firms' exit decisions are determined by their perceptions of future productivity and these perceptions partially depend on their current productivity, then a balanced panel will generate a selection bias. Since low productivity firms are more likely to exit the sample than their more productive counterparts, omitting them from the analysis would lead to an upward bias in estimated productivity.

A relatively easy method in response to the aforementioned methodological problems is assuming that the part of TFP that influences firm behaviour is a firm-specific attribute and it doesn't vary over time. If the assumption holds, estimating a fixed-effect panel regression will solve the simultaneity problem; however, different studies showed that fixed-effects estimators have not proved satisfactory for the case of production functions (Arnold, 2005). Another alternative to estimate production function parameters consists in using instrumental variables, but the difficulty in this case arises from finding appropriate instruments. Olley and Pakes (OP, 1996) were the first to introduce a semi-parametric estimator to deal with both endogeneity and selection problem. Their estimator deals with simultaneity using firm's investment decision as a proxy for productivity shocks, while the selection problem is solved by generating a survival rule. However, because many firms do not report non-zero investment levels, Levinsohn and Petrin (2003) have proposed that intermediate inputs should be used as a proxy for productivity. Their algorithm does not include

[3]Other problems related to omitted output price bias, omitted input price bias or multi-product firms have been addressed in the literature (see Van Beveren, 2008).

a survival probability, as the efficiency gains compared to an unbalanced panel are found to be very small.

3.2. Estimation

In order to estimate TFP in a consistent manner and to document the productivity differentials between exporters and non-exporters, we estimate production functions for each of the ten business sectors, starting from a Cobb-Douglas technology:

$$y_{it} = \beta_0 + \beta_k k_{it} + \beta_l l_{it} + \omega_{it} \qquad (1)$$

where y_{it} is the log of output (value added) from firms at time t, $k_{it}$the log of capital input, l_{it} the log of labourβ_0 the sector specific intercept,$\beta_l$and β_k are the output elasticities of labour and capital respectively and ω_{it} is the residual term, which is used for estimating productivity, $t = \overline{2004, 2011}$.

We estimate eq. (1) by means of OLS. However, as already mentioned, this method is invalidated because the endogeneity bias. In consequence, we also employed a second method, namely fixed effects panel regression.

$$y_{it} = \beta_0 + \beta_k k_{it} + \beta_l l_{it} + \omega_i + \eta_{it} \qquad (2)$$

where y_{it}, k_{it}, l_{it}, β_0,$\beta_l$and β_k are defined above, ω_i is firm-specific but time-invariant (fixed effect4) and η_{it} is a i.i.d. remainder disturbance, that varies over time and entities and captures everything that is left unexplained about y_{it}.In all the estimates robust standard errors are computed, as error testing indicates cross-section heteroscedasticity (Modified Wald test for group wise heteroscedasticity in fixed effect regression model) and no serial correlation at a 5% significance level (Wooldridge test for autocorrelation in panel data).

Using the semi-parametric approach suggested by Levinsohn and Petrin (2003), the firm's output is written as:

$$y_{it} = \beta_0 + \beta_k k_{it} + \beta_l l_{it} + \beta_m m_{it} + \omega_{it} + \eta_{it} \qquad (3)$$

where m_{it} is the log level of materials, expressed as function (which can be inverted) of capital and productivity and which is used as a proxy for TFP:

$$m_{it} = m_{it}(\omega_{it}, k_{it}) \qquad (4)$$

$$\omega_{it} = \omega_{it}(m_{it}, k_{it}) \qquad (5)$$

The productivity index $\omega_{it}$is known to the firm and evolves over time according to an exogenous Markov process:

$$\omega_{it} = E[\omega_{it} | \omega_{it-1}] + \xi_{it} \qquad (6)$$

where $\xi_{it}$is productivity innovation.

The production function can be written as:

$$y_{it} = \beta_l l_{it} + \Phi(m_{it}, k_{it}) + \eta_{it} \qquad (7)$$

[4]The fixed effects for the model were confirmed by the Hausman test (1978).

$$\Phi(m_{it}, k_{it}) = \beta_0 + \beta_k k_{it} + \omega_{it}(m_{it}, k_{it}) \quad (8)$$

In the first step of the algorithm, function $\Phi(m_{it}, k_{it})$ is approximated with a 3rd order polynomial in (m_{it}, k_{it}) and estimates for β_l is obtained. In the second stage, β_k is identified and all these parameters are used for predicting a non-parametric approximation for $E[\omega_{it}|\omega_{it-1}]$. In the final step, the residual of the production function is obtained. In value added case, TFP is expressed as:

$$TFP_{it} = exp(y_{it} - \widehat{\beta}_k k_{it} - \widehat{\beta}_l l_{it}) \quad (9)$$

The last method of estimating TFP is the one step procedure proposed by Wooldridge (2009), in which standard GMM is used. The error term is assumed to be uncorrelated with labour and capital and also with lags of these variables:

$$E[\eta_{it}|k_{it}, l_{it}, m_{it}, k_{it-1}, l_{it-1}, m_{it-1}, \ldots, k_{i1}, l_{i1}, m_{i1}] = 0\,, \quad t=1,2,\ldots,T \quad (10)$$

where m_{it} is in this case the intermediate inputs.

Another assumption is to restrict the dynamics of unobserved productivity process:

$$E[\omega_{it}|k_{it}, k_{it-1}, l_{it-1}, m_{it-1}, \ldots, k_{i1}, l_{i1}, m_{i1}] = E[\omega_{it}|\omega_{it-1}] \equiv f[g(k_{it-1}, m_{it-1})] \quad (11)$$

Productivity innovations are allowed to be correlated with variable inputs (l_{it}, m_{it}), while these are uncorrelated while the state variable (k_{it}) and all past values of k_{it}, l_{it} and m_{it}. As contemporaneous capital variable, lagged inputs and function of these can be used as instrumental variables, we use as instruments for l_{it}, $k_{it}, k_{it-1}, l_{it-1}, m_{it-1}$ and up to three order polynomials containing capital and materials.

In order to obtain comparable TFP levels across firms from different sectors, it is expressed as an index, by comparing each firm to the average over all firms in the respective sector in a certain year:

$$\text{(Relative) } TFP_{it} = \ddot{TFP}_{it} / \overline{TFP} \quad (12)$$

where $\ddot{TFP}_{it}$ is the TFP (estimated as in equation 9), for a firm i at moment t and $\overline{TFP}$ is the sector specific productivity (average at moment t).

Table 1. Production function coefficients (2004-2011)

Sector	Production function coeff.	*Levinsohn and Petrin*	*OLS*	*Fixed Effects*	*Wooldridge*
Mining	labour	*0.464	*0.779	*0.694	*0.708
		(0.0296)	(0.0183)	(0.0432)	(0.0338)
	capital	*0.372	*0.292	*0.249	*0.373
		(0.0350)	(0.0128)	(0.0265)	(0.0320)
Food, beverage and tobacco	labour	*0.460	*0.870	*0.674	*0.842
		(0.0097)	(0.0060)	(0.0141)	(0.0115)
	capital	*0.157	*0.291	*0.250	*0.177
		(0.0100)	(0.0037)	(0.0083)	(0.0105)

Textile, leather	labour	*0.704 (0.0072)	*0.818 (0.0047)	*0.585 (0.0121)	*0.837 (0.0097)
	capital	*0.161 (0.0084)	*0.244 (0.0034)	*0.168 (0.0075)	*0.187 (0.0092)
Chemicals, pharmaceutical products	labour	*0.473 (0.0030)	*0.901 (0.0171)	*0.719 (0.0404)	*0.873 (0.0312)
	capital	*0.188 (0.0250)	*0.233 (0.0105)	*0.184 (0.0206)	*0.169 (0.0266)
Basic metal/Metallurgic industry	labour	*0.636 (0.0099)	*0.826 (0.0064)	*0.663 (0.0152)	*0.864 (0.0125)
	capital	*0.220 (0.0123)	*0.265 (0.0041)	*0.225 (0.0088)	*0.224 (0.0104)
Electrical and electronic products	labour	*0.636 (0.0207)	*0.825 (0.0106)	*0.654 (0.0328)	*0.846 (0.0224)
	capital	*0.169 (0.0179)	*0.230 (0.0077)	*0.174 (0.0159)	*0.179 (0.0191)
Auto industry	labour	*0.653 (0.0165)	*0.823 (0.0090)	*0.682 (0.0250)	*0.831 (0.0197)
	capital	*0.160 (0.0157)	*0.205 (0.0063)	*0.174 (0.0132)	*0.169 (0.0158)
Electricity	labour	*0.258 (0.0505)	*0.665 (0.0224)	*0.646 (0.0744)	*0.635 (0.0468)
	capital	*0.315 (0.0446)	*0.264 (0.0166)	*0.267 (0.0427)	*0.282 (0.0522)
Construction	labour	*0.535 (0.0049)	*0.821 (0.0028)	*0.733 (0.0060)	*0.820 (0.0058)
	capital	*0.264 (0.0052)	*0.308 (0.0018)	*0.179 (0.0034)	*0.261 (0.0043)
Telecommunications and other services	labour	*0.625 (0.0272)	*0.840 (0.0119)	*0.672 (0.0221)	*0.827 (0.0304)
	capital	*0.210 (0.0158)	*0.301 (0.0074)	*0.178 (0.0121)	*0.222 (0.0159)

Robust standard errors in parentheses. * significant at 1% level; ** significant at 5%; *** significant at 10%.

Table 1 shows production function coefficients estimated using OLS, fixed effects, Levinsohn and Petrin (2003)and Wooldridge (2009)methodology. All regressions yield reasonable coefficients for production functions' parameters, production behaviour varying between sectors (as input combinations and demand elasticity differ and labour markets are not homogeneous). As expected, the OLS coefficients are much higher, the semi-parametric approach yielding to lower coefficients, especially for labour. Despite these differences, the TFP distribution seems not to be extremely sensitive to the choice of estimation procedure. In what follows, results from Levinsohn and Petrin (2003) approach are used.

The resulting shapes of TFP distribution, estimated for each business sector, are asymmetric, displaying fat tails to the right (Figure 1). This result is in line with other studies focusing on European countries (e.g. Lopez-Garcia et al., 2015; CompNet, 2014; Mayer and Ottaviano, 2007). This means that there are large heterogeneities not only between sectors, but also within the same sector, i.e. there is a large number of relatively “bad” firms having a TFP index below the mean, but also a certain number of particularly good firms, thus determining a distribution with a relatively long right tail.

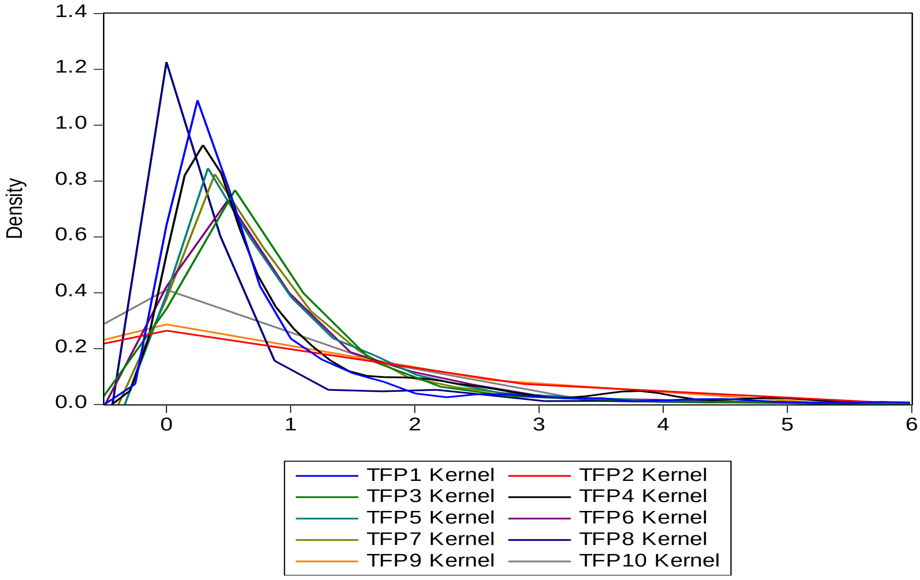

Figure 1. Kernel density estimation for TFP across business sectors (2011)
Note: Distributions of firms’ TFP in each sector, relative to sector’s average

3.3. TFP Analysis. Traders versus non-traders, a primer

After obtaining estimates of the production function and subsequently for TFP (equation 12) the link between exporting activities and productivity is investigated. The TFP distributions for exporters and domestic only firms are ranked using the concept of first order stochastic dominance, and their differences are tested using Kolmogorov–Smirnov one and two-sided tests (Kolmogorov, 1933 and Smirnov, 1939). More formally, if F and G are the cumulative distribution functions of traders and non-traders’ productivity, then first order stochastic dominance of F relative to G is written as: $\mathrm{F(x)} - \mathrm{G(x)} \leq 0$ for all $x \epsilon$ R, with strict inequality for some x.

(i) Two sided test:

$$H_0: F(x) - G(x) = 0 \text{ for all } x \in R \quad (13)$$

If H_0 is rejected, the two distributions are not drawn from the same underlying continuous distribution and it makes sense to employ the one-sided test, which allows to determine whether a distribution dominates the other (if H_0 cannot be rejected).

(ii) One-sided test:

$$H_0: F(x) - G(x) \leq 0 \text{ for all } x \in R \quad (14)$$

Table 3. Kolmogorov-Smirnov test for testing differences in TFP distributions between exporters/importers and non-traders

Sector	One sided test		Two sided test (*p*-value)	One sided test		Two sided test (*p*-value)
	All exporters	All non-traders		**All importers**	All non-traders	
Mining		-		-0.008	*0.865	0
Food, beverage and tobacco	-0.002	*0.779	0	0.000	*0.868	0
Textile, leather	0.000	*0.558	0	-0.004	*0.463	0
Chemicals, pharmaceutical products	0.000	*0.687	0	-0.002	*0.646	0
Basic metal/Metallurgic industry	-0.003	*0.542	0	-0.009	*0.574	0
Electrical and electronic products	-0.001	*0.454	0	0.000	*0.541	0
Auto industry	-0.001	*0.473	0	-0.006	*0.501	0
Electricity	0.000	*0.778	0.005	-0.035	**0.536	0.009
Construction	-0.014	*0.597	0.001	0.000	*0.713	0
Telecommunications and other services		-		-0.009	*0.750	0

Note: * denotes null rejected at 1% level; ** null rejected at 5% level; *** null rejected at 10% level.

Romanian data shows that exporting companies account for over one third of total imports' volume (as of 2011), almost 70% of exporters having significant importing activity (over EUR 100,000 in each quarter). Thus, following Altomonte and Bekes (2009) who revealed the importance of importing activities in analysing the link between involving export and productivity, we also investigate the behaviour of importers. Figure 2 highlights important differences between the productivity levels of trading and domestic only companies. Moreover, the TFP distribution for exporters and importers stochastically dominates the non-traders' (Table 3).

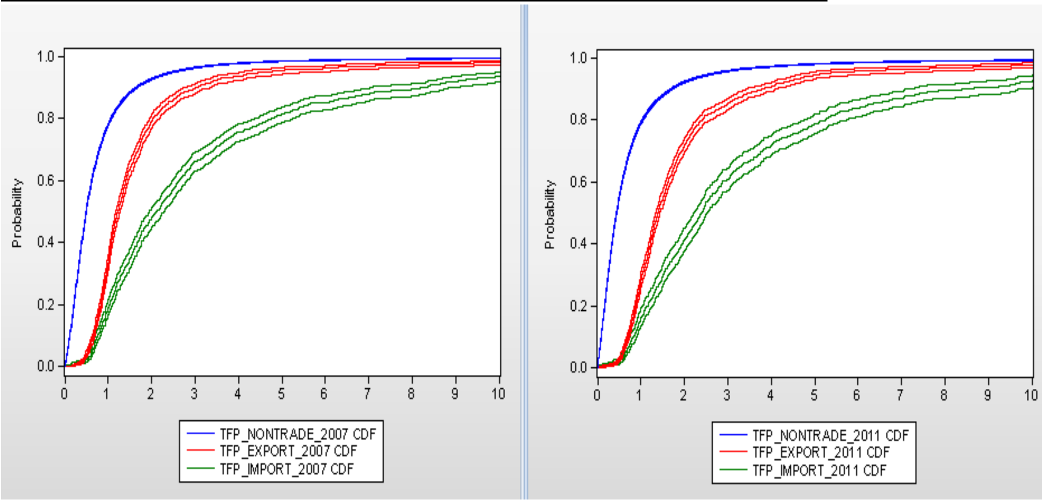

Figure 2. Empirical cumulative distribution functions of TFP. Traders versus non-traders (2007, left side; 2011, right side)

4. Methodology for testing export-productivity link

4.1. Propensity score matching

One of the most important methodological problems emerged when testing the effect of exporting on firm-level performance is linked to the sample-selection bias. This arises when making comparison between the group of trading firms (the treatment group, the treatment being export status) and the rest of the companies. As the firms in the treatment group have, most probably, managed to become exporters due to some unobservable characteristics estimating learning effects using traditional econometric routines would lead to biased results. Matching methods are used as an efficient instrument to deal with problems arising from endogenous participation decision. These methods rely on building a suitable control group from among non-traders that will be used as counterfactuals for exporters. This control group should have$n-1$ (out of n) features similar to the exporters group and differ only in the nthcharacteristic, which is the decision to export.

For assessing the causal impacts of exporting on productivity, we employ propensity score matching method to select the control group, following Rosenbaum and Rubin (1983). Propensity score matching corrects the assessment of the effects of the treatment, controlling for the existence of confounding factors5 and reduces the bias by making comparison of outcomes, using treated and control subjects who are as similar as possible. Matching on an n-dimensional vector of characteristics is typically unfeasible for large n, leading to the problem known as

[5]correlated with both the dependent variable and the independent variables.

the curse of dimensionality. The propensity score summarizes all relevant pre-treatment characteristics of each subject into a single-index variable and thus makes the matching feasible, the bias being eliminated only if the exposure to treatment can be considered to be purely random among individuals who have the same value of the propensity score. For the scope of this paper, Propensity score matching is used in order to detect some non-exporting firms that had similar tendency to export as exporting companies but in fact stayed only on the domestic markets. The main element of interest, as in all matching methods, is the Average Treatment effect on the Treated (ATT), which means the difference for each "treated" firm between: (i) the effective outcome it obtains as an exporter (the treatment is export activity) and (ii) the potential outcome it would have obtained if it had chosen not to involve in export sales.

$$\text{ATT}=E\left[Y_{i,t}(1)-Y_{i,t}(0)\middle|Export_{i,t}=1\right]= E\left[Y_{i,t}(1)\middle|Export_{i,t}=1\right]-E\left[Y_{i,t}(0)\middle|Export_{i,t}=1\right] \quad (15)$$

where: $Y_{i,t}(1)$is the outcome of a firm iin year t ($t=\overline{2007,2011}$)given it is an exporter in year t, as defined is the Section2; As outcomes, we use TFP (estimated as in Section 3), as well as Labour productivity, measured as real value added per unit of employee, and capital productivity, measured as real value added per unit of real capital;$Y_{i,t}(0)$the outcome of the firm if it didn't export in the specific year; $Export_{i,t}$takes value 1 if the firm i decided to export at moment t. Since$E\left[Y_{i,t}(0)\middle|Export_{i,t}=1\right]$ cannot be observed from the data, it is referred to as the counter factual outcome. However, we can compute the outcome for non-exporters provided that they have not exported i.e. $E\left[Y_{i,t}(0)\middle|Export_{i,t}=0\right]$, which is used to replace the unobservable outcome of exporting firms if they hadn't been exporters. From this replacement, if export entry is non-random,a bias equal to $E\left[Y_{i,t}(0)\middle|Export_{i,t}=1\right]$ - $E\left[Y_{i,t}(0)\middle|Export_{i,t}=0\right]$can arise. In consequence, some assumptions must be made in order to eliminate the selection bias. The conditional independence assumption states that the variables on which the matching is done are not affected by the treatment, either ex-post or in anticipation of the treatment. Also, conditional on the set of covariates $Z_{i,t-1}$, the outcome Y is independent of the export decision:

$$Y_{i,t}(1),\ Y_{i,t}(0) \perp \text{Export} \mid Z_{i,t-1} \quad (16)$$

When the conditional independence assumption holds, one can use the productivity of firms not exporting as an approximation of the counterfactual outcome. Heckman et al. (1998) show that for an unbiased estimation of ATT, it is only necessary to assume:

$$E\left[Y_{i,t}(0)\middle|Z_{i,t-1},\ Export_{i,t}=1\right]=E\left[Y_{i,t}(0)\middle|Z_{i,t-1},\ Export_{i,t}=0\right] \quad (17)$$

Another assumption required to ensure that the counterfactual groups can be created is that$Z_{i,t-1}$is not a perfect predictor of treatment status. This ensures that for every $Z_{i,t-1}$, there are firms choosing to export and firms choosing not to export. The common support condition is imposed and the balancing hypothesis is ensured (firms with the same propensity score must have the same distribution of observable and unobservable features independently of treatment status).

We assume that actual export choice can be described by a latent variable whose distribution determines the export decision:

$$Export_{it}^{*} = h(Z_{it-1}) + \varepsilon_{it} \quad (18)$$

where$Export_{it}^{*}$ is the latent variable, $h(Z_{it-1})$ is a function of firm observable characteristics and $\varepsilon_{it}$is the error term, assumed to be normally distributed and homoscedastic.

If $Export_{it}^{*}$ exceeds a certain threshold, then the company exports:

$$Export_{it} = 1\ if\ Export_{it}^{*} > 0\ and\ 0\ otherwise \quad (19)$$

Thus, using a standard normal distribution, the probability of exporting can be formulated as a probit regression:

$$\mathrm{P}(Export_{it} = 1|Z_{it-1}) = \mathrm{P}(Export_{it}^{*} > 0\ |Z_{it-1}) = \Phi(h(Z_{it-1})) \quad (20)$$

Namely, the probability for a firm to be an exporter in the year t, t from 2008 to 2011, can be modelled as a cumulative distribution function $\Phi(h(.))$, where $h(.)$ is a polynomial function of the covariates. We use as covariates in estimating the propensity score6 lagged TFP, lagged size (measured as the logarithm of the number of employees) and a dummy indicating the ownership of the firm. Other variables have been tested (e.g. return on equity, leverage) but were dropped as these affected the balancing property or were statistically insignificant. The score and the matching are conducted separately for each business sector.

The results indicate that larger companies (in terms of employment) are more likely to engage in export activities, foreign ownership having also a positive impact on export decision. Most importantly, TFP in the previous year has positive impacts on the likelihood of being an exporter in the current year (increasing the propensity score), indicating the presence of the self-selection process into exporting market. The results are statistically significant in most of the cases (Table 4). Possible explanations why only the most productive firms self-select into export sales can derive from barriers and initial costs related to exporting activities (e.g. costs of establishing new distribution/ marketing channels/ market research, the need to establish contacts in the destination country, support centres, product modifications) that can be overcome only by the performing companies.

[6]These are in line with a number of studies analyzing the determinants of firm's decision to export e.g. De Loecker (2007).

Table 4.Probit regression results. Exporters

Sector	TFP t-1	number of employees t-1	ownership	Pseudo R-squared
Food, beverage tobacco	*0.346 (0.0257)	*0.002 (0.0003)	*0.710 (0.0660)	0.338
Textile, leather	*0.136 (0.0063)	*0.006 (0.0001)	*0.747 (0.0221)	0.313
Chemicals, pharma products	*0.576 (0.0667)	*0.002 (0.0002)	*0.7166 (0.0859)	0.309
Basic metal	*0.191 (0.0171)	*0.003 (0.0001)	*1.194 (0.0332)	0.294
Electrical products	*0.0736 (0.0209)	*0.004 (0.0003)	*1.373 (0.0645)	0.392
Auto industry	*0.130 (0.0168)	*0.001 (0.0001)	*0.979 (0.0439)	0.281
Electricity	*0.154 (0.0260)	0.001 (0.0008)	*0.441 (0.1379)	0.209
Construction	**0.066 (0.0270)	*0.002 (0.2079)	*0.517 (0.1034)	0.199

Robust s.e. in parentheses.*significant at 1%;** significant at 5%;*** significant at 10%

By estimating a similar probit model for explaining the propensity to import (Table 5), positive link between lagged TFP and probability of being an importer in current period is found, which indicate that self-selection of most productive firms into importing activities also holds in Romanian case.

Table 5.Probit regression results. Importers

Sector	TFP t-1	number of employees t-1	ownership	Pseudo R-squared
Mining	*0.373 (0.0594)	*0.001 (0.0001)	*1.278 (0.1358)	0.397
Food, beverage and tobacco	*0.0878 (0.0116)	*0.005 (0.0003)	*0.403 (0.0509)	0.365
Textile, leather	*0.223 (0.0346)	*0.005 (0.0004)	*0.534 (0.0452)	0.195
Chemicals, pharmal products	*0.619 (0.0379)	*0.001 (0.0003)	*0.535 (0.0769)	0.352
Basic metal	*0.584 (0.0683)	*0.003 (0.0002)	*0.724 (0.0483)	0.249
Electrical products	*0.288 (0.0247)	*0.003 (0.0003)	*0.561 (0.0697)	0.247
Auto industry	*0.839	*0.002	*0.730	0.274

	(0.0745)	(0.0002)	(0.0514)	
Electricity	*0.326	*0.001	*0.378	0.232
	(0.0591)	(0.0001)	(0.1364)	
Construction	*0.049	*0.002	*0.689	0.172
	(0.0077)	(0.0001)	(0.0497)	
Telecommunications and other services	*0.406 (0.0363)	*0.001 (0.0002)	**0.251 (0.0999)	0.327

Robust s.e. in parentheses.*significant at 1%;** significant at 5%;*** significant at 10%

4.2. Nearest neighbour and Kernel matching

After estimating the propensity scores for each sector, we pair exporters (treated firms) and controls (non-treated firms). However, the estimate of the propensity score $p(X)$ is not enough to estimate the ATT of interest since the probability of observing two firms with exactly the same value of the propensity score theoretically zero, $p(X)$ being a continuous variable. For pairing the firms, we applied nearest-neighbour matching (with replacement) and kernel matching. The former method takes each exporter (treated unit) and pairs it with a single non-exporter (control unit) by minimizing the absolute difference between the estimated propensity scores for the control and treatment groups. The kernel method, matches all treated units with a weighted average of all controls, using weights that are inversely proportional to the distance between the propensity scores of exporters and non-exporters.

Once each exporter is matched with a control firm, we computed the difference between the outcome of the exporters and the outcome of the matched non-exporters. In this case, we define the outcome variables as the level of TFP relative to sectors' average and the (logarithmic) level of labour productivity and capital productivity. The ATT of interest is then obtained by averaging the differences between the two matched groups. In the following, the kernel matching results are reported, the nearest-neighbour method revealing in most of the cases similar conclusions. In estimating the variance of the treatment effect, we applied bootstrapping7 method. The procedure is replicated for importers as well, as estimating the impact of importing activity of firm's performance is of particular interest, as already mentioned.

4.3. Matching results

The results suggest that the effect of exporting on TFP is generally positive, but it is found statistically significant only in three sectors (Table 6). The results also suggest that taking into account only labour productivity and not the

[7]For each estimation, a new sample of the same size is drawn with replacement and all the steps of the estimation, including simulation of start dates and the enforcement of common support, are performed on the simulated sample.

multi-factorial productivity would lead to an overstated exporters' advantage (as exporters tend to have an unfavourable position regarding capital productivity).Somehow surprising is finding significant productivity differential between exporters and non-exporters in low/medium low-tech industries. Nevertheless, learning-by-exporting can still be possible for firms in these sectors as these can still benefit from their contacts abroad to get technological information or to gain from being in a higher competition environment.

Table 6. Kernel matching results. Exporters

Sector	TFP		Labour productivity		Capital productivity	
	ATT	Bootstrap s.e.	ATT	Bootstrap s.e.	ATT	Bootstrap s.e.
Food, beverage, tobacco	*1.966	(0.340)	*1.157	(0.087)	*0.176	(0.073)
Textile, leather	*0.190	(0.072)	*0.389	(0.037)	*0.490	(0.032)
Chemicals, pharma	0.162	(0.126)	*0.438	(0.057)	**-0.192	(0.107)
Basic metal/Metallurgic industry	0.072	(0.061)	*0.393	(0.035)	*-0.255	(0.057)
Electrical and electronical products	0.020	(0.126)	*0.355	(0.038)	*0.245	(0.060)
Auto industry	0.060	(0.129)	*0.329	(0.038)	*-0.248	(0.043)
Electricity	3.288	(2.238)	*1.430	(0.366)	**1.866	(0.905)
Construction	**1.902	(0.783)	*0.864	(0.178)	0.008	(0.168)

Note: *significant at 1% level; ** significant at 5%; *** significant at 10%

Table 7. Kernel matching results. Importers

Sector	TFP		Labour productivity		Capital productivity	
	ATT	Bootstrap s.e.	ATT	Bootstrap s.e.	ATT	Bootstrap s.e.
Mining	*2.002	(0.480)	*1.147	(0.162)	0.234	(0.166)
Food, beverage and tobacco	**2.478	(1.051)	*1.015	(0.034)	*0.104	(0.035)
Textile, leather	*0.395	(0.088)	*0.561	(0.027)	*0.217	(0.074)
Chemicals, pharmaceutical products	*0.426	(0.131)	*0.492	(0.072)	*-0.297	(0.089)
Basic metal/Metallurgic industry	*0.476	(0.082)	*0.702	(0.054)	*-0.749	(0.040)
Electrical and electronic products	**0.722	(0.278)	*0.565	(0.089)	**-0.183	(0.092)
Auto industry	*0.144	(0.052)	*0.423	(0.043)	*-0.469	(0.071)
Electricity	**0.668	(0.313)	*0.583	(0.196)	0.067	(0.306)
Construction	*3.141	(0.447)	*1.667	(0.093)	*0.408	(0.088)
Telecommunications and other services	*0.864	(0.267)	*0.920	(0.152)	**-0.250	(0.142)

Note: *significant at 1% level; ** significant at 5%; *** significant at 10%

In importing companies' case however, the positive productivity differential is found significant in most cases (Table 7), suggesting that for Romanian case learning-by-importing hypothesis might hold. The relatively higher premium for importers manifested in capital and intermediate goods, can be explained by the fact that thorough importing these goods (that cannot be produced domestically/are of inferior quality in domestic markets) firms are enabled to diversify, specialize and further enhance their productivity. These imports can permit upgrading own products and create new export possibilities.

5. Concluding remarks

In this paper, we examine the causal relationship between exporting and productivity at the firm level, the main indicator used being unobservable TFP, estimated byLevinsohn and Petrin(2003). Using business sectors models, our results showed that Romanian exporters seem to be more productive than non-exporters. We found that productivity increases the probability of exporting in most sectors, confirming self-selection of most productive firms in exporting activities. In order to determine the effect of exporting to productivity and test learning-by-exporting hypothesis, we use matching techniques to control for the non-random selection of exporting firms in the sample. The productivity differences between exporting and non-exporting firms within the matched pairs are significant only on a narrow sample of companies, which suggest that learning effects are highly heterogeneous. Thus, the results suggest that the direction of causality runs from productivity to exporting while the direction of causality from exporting to productivity is less clear. However, since importing and exporting activities are highly correlated within firms, we show that importing companies exhibit higher productivity differential. The result suggests that, without controlling for importing activity, an overstated average productivity premium of exporters can occur. Also, self-selection of performing firms in importing is confirmed and the learning effects seem to be systematically related to both importing and exporting activities. One possible explanation is that the more complex the trade activities a firm is involved in, the more productive it needs to be.

Potential policy implications revealed form this study are that (1) exporting is not a panacea for enhancing productivity, since firms must have high productivity before they can involve in external market sales and (2) although exports' reliance on cost of imported inputs can represent a challenge for Romanian exports, when formulating policies in order to promote/discourage importing activities, one has to take into account the import purpose, variety and quality (as superior imported inputs can lead to improved productivity and new export opportunities). Further research areas could be the analysis of export destinations' impact on the magnitude of learning effects, the productivity differential of companies that benefit from FDI and a more in depth analysis of import and productivity link.

ACKNOWLEDGMENT:

This work was supported by a grant of the Romanian National Authority for Scientific Research, CNCS – UEFISCDI, project number PN-II-ID-PCE-2011-3-1054.

REFERENCES

[1] **Altăr, M.,Necula, C. and Bobeică, G. (2010),** ***Estimating Potential GDP for the Romanian Economy. An Eclectic Approach****; Romanian Journal of Economic Forecasting*, 3/2010, pp. 5-25;

[2] **Altomonte, C. and Bekes, G. (2009),** ***Trade Complexity and Productivity***; *IEHAS Discussion Papers 0914*;

[3] **Arnold, J.M. (2005),** ***Productivity Estimation at the Plant Level: A Practical Guide***; mimeo, Bocconi University;

[4] **Bartelsman, E., Haltiwanger, J. and Scarpetta, S. (2013),** ***Cross-Country Differences in Productivity: The Role of Allocation and Selection*****;** *American Economic Review*; American Economic Association, vol. 103(1), pages 305-34;

[5] **Berman, N., Berthou, A. and Hericourt, J. (2015),** ***Export Dynamics and Sales at Home;*** *Journal of International Economics*, vol. 96 (2), pages 298-310;

[6] Bernard, A.B., Wagner, J. (1997), *Exports and Success in German Manufacturing; Review of World Economics*, 133(1), 134-157;

[7] **Berthou, A., Dhyne, E., Bugamelli, M., Cazacu A.M., Demian, V., Harasztosi, P., Lalinsky, T., Meriküll, J., Oropallo, F. and Soares, A. (2015),** ***Assessing European Firms' exports and Productivity Distributions: The CompNet Trade Module****; ECB Working Paper Series,* no. 1788;

[8] **Cazacu, A.M. (2014),** ***Foreign Trade - Productivity Nexus. What Can Firm-level Data Tell us?*****;** *LAP Lambert Academic Publishing,* ISBN 3659636053;

[9] **Competitiveness Research Network (2014),** ***Micro-based Evidence of EU Competitiveness. The CompNet Database***; *ECB Working Paper* 1634;

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

[11] **De Loecker, J (2013),** ***Detecting Learning by Exporting***; *American Economic Journal: Microeconomics*, 5: 1-21;

[12] **Hagemejer, J. (2007),** ***Factors Driving the Firm's decision to Export. Firm-level Evidence from Poland***; *MPRA Paper* No. 17717;

[13] **Hansen, D.M. and Smith, V.(2015), R&D, *Export and Productivity in Business Services Firms: Testing the Bustos Model***; *Applied Economics Letters*, 22(16);
[14] **Heckman, J.J., Ichimura, H. and Todd,P.E. (1998), *Matching as an Econometric Evaluation Estimator***; *Review of Economic Studies*, 65, 261-294;
[15] **Hosono, K., Miyakawa, D. and Takizawa, M. (2015), *Learning by Export: Does the Presence of Foreign Affiliate Companies Matter***? *RIETI Discussion Paper*, 15-E-053;
[16] **Levinsohn, J. and Petrin, A. (2003), *Estimating Production Function Using Inputs to Control for Unobservable***; *Review of Economic Studies* 70, 317–341;
[17] **Lopez, R. A. (2004),*Self-Selection into Export Markets: A Conscious Decision?";*** mimeo, *Department of Economics, Indiana University*;
[18] **Lopez-Garcia, P., di Mauro, F. and CompNet Task Force (2015) , *Assessing European Competitiveness: The New CompNet Micro-based Database***; *ECB Working Paper Series*, no. 1764;
[19] **Mayer, T. and Ottaviano, G.I.P. (2007), *The Happy Few: The Internationalization of European Firms***; *Bruegel Blueprint Series*;
[20] **National Bank of Romania (2010),*Financial Stability Report***,[online] Available at: <http://bnr.ro/files/d/Pubs_en/RSF/eRSF2010.pdf>[Accessed in January 2015];
[21] **Olley, G.S. and Pakes, A. (1996),*The Dynamic of Productivity in the Telecommunication Industry***; *Econometrica* 64(6), 1263–97;
[22] **Pisu, M. (2008),*Export Destinations and Learning-by-Exporting: Evidence from Belgium***; *National Bank of Belgium, Working Papers* No. 140;
[23] **Rosenbaum, P.R. and Rubin, D.B. (1983),*The Central Role of the Propensity Score in Observational Studies for Causal Effects***; *Biometrika* 70, 41–55;
[24] **Solow, R.M. (1957),*Technical Change and the Aggregate Production Function***; *The Review of Economics and Statistics*, Vol. 39, No. 3., pp. 312-320;
[25] **Wooldridge, J.M. (2009),*On Estimating Firm-level Production Functions Using Proxy Variables to Control for Unobservables***; *Economic Letters*, 104.