

Jienwatcharamongkhol, Viroj; Tavassoli, Mohammad H.

Conference Paper

Closing the Gap: An Empirical Evidence on Firm's Innovation, Productivity, and Export

53rd Congress of the European Regional Science Association: "Regional Integration: Europe, the Mediterranean and the World Economy", 27-31 August 2013, Palermo, Italy

Provided in Cooperation with:

European Regional Science Association (ERSA)

Suggested Citation: Jienwatcharamongkhol, Viroj; Tavassoli, Mohammad H. (2013) : Closing the Gap: An Empirical Evidence on Firm's Innovation, Productivity, and Export, 53rd Congress of the European Regional Science Association: "Regional Integration: Europe, the Mediterranean and the World Economy", 27-31 August 2013, Palermo, Italy, European Regional Science Association (ERSA), Louvain-la-Neuve

This Version is available at:

<https://hdl.handle.net/10419/123956>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Closing the Gap: An Empirical Evidence on Firm's Innovation, Productivity, and Export

Viroj Jienwatcharamongkhon

Department of Economics, Lund University, Sweden

Mohammad H. Tavassoli

School of Management, Blekinge Institute of Technology, Sweden

Abstract

It is well known that exporters are productive firms. But the source of their productivity is left unexplained. This paper aims to endogenize the productivity heterogeneity of exporting firms by incorporating innovation in a structural model framework. In doing so, we close the gap between the innovation-productivity and productivity-export literature. Two waves of Swedish Community Innovation Survey (CIS) are merged. This allows for a dynamic setup that takes into account the proper lags from innovation input to innovation output and also from innovation output to productivity and export. The applied framework corrects for selectivity, simultaneity, and potential endogeneity biases. The main findings highlight that exporters are productive firms with innovation output in the past, which in turn was driven by prior R&D and other innovation activity investments.

Keywords: innovation, productivity, export, firm-level, structural model, community innovation survey

JEL: C31, L60, O31

Acknowledgement: Financial support from Torsten Söderberg Foundation is gratefully acknowledged.

1. Introduction

Exporters are known to be a selected group of productive firms. They can afford the associated upfront fixed costs of entering the foreign markets. In other words, they succeed in crossing the *productivity threshold* and *self-select* themselves into export (Bernard & Jensen, 1999; Bernard & Wagner, 1997; Delgado, Fariñas, & Ruano, 2002; Melitz, 2003). This self-selection literature has been an important theoretical foundation for many of the recent empirical studies among trade scholars, mainly due to its prediction that is in line with the observed data. Besides being more productive, exporters are a bigger-sized minority among firms – a stylized fact that Eaton, Kortum, and Kramarz (2004) observe for French manufacturing firms¹. However, the literature is still in its infancy and the existing evidence often provides a mixed picture. For example, the direction of causation between productivity and export is still an unresolved debate. The self-selection literature treats firm's productivity as exogenous and thus it remains a “black box.” Where does the heterogeneity of firm's productivity come from?

There are attempts to formerly model the endogenous firm productivity. In Ederington and McCalman (2008), the defining element in their model is the adoption rate of new technology or innovation. Another model by Segerstrom and Stepanok (2011) involves two types of firm's R&D investments – traditional quality improvement activities and investments for becoming exporters. Similarly, if we look at another strand of literature, namely micro-studies of innovation, it is found that the main source of productivity heterogeneity is the firm's investment in research and development (R&D) and other innovation activities that lead to innovation output of a firm². The production function framework, advocated by Griliches (2000), and recently the endogenous growth model (P. Aghion & Howitt, 1992; Romer, 1990) explain productivity gain as a result of capital accumulation and technological change (innovation). Despite the innovation being a possible answer to the question we posed above, the study that attempts to examine this is still rare³.

This paper examines the link from R&D investments and innovation output to the productivity and export performance. In doing so, we close the gap between the self-selection and innovation-productivity literature. The empirical evidence in this paper comes from the modified structural model that extends the framework by Crépon, Duguet, and Mairesse (1998) and (Löf & Heshmati, 2006). Empirically, we employ two waves of Community Innovation Survey (CIS) done for Sweden and complement them with detailed firm export and registry data. The panel setup gives us the flexibility to interact the innovation, productivity, and firm's export. Moreover, we are able to deal with several econometric issues, namely the selectivity, simultaneity, and endogeneity problems.

¹ Other studies have found the same pattern. See a review by Wagner (2007) for such studies.

² The recent studies are based on innovation survey data. For a review, see Hall and Mairesse (2006).

³ Antonietti & Cainelli (2011) is an exception.

Because not all firms invest in R&D and other innovation activities, excluding the non-innovative firms from the estimation will give rise to the selectivity problem. While allowing the interaction between innovation, productivity, and export, we are able to deal with the simultaneity issues, since it is argued that export can also raise the firm's incentives to innovate (Long, Raff, & Stähler, 2011). The dynamic nature of innovation process involves a lag time, thus disregarding it will result in the endogeneity problems. We will discuss more in details later.

Our main contribution is the following. First, this paper adds to the discussion of firm's heterogeneity by endogenizing it with innovation activities. Second, we extend the existing structural model by including firm's export in order to examine the relationship between innovation, productivity, and export, providing support for our argument. Third, the panel structure allows us to properly address the dynamics of innovation process and the econometric issues, which are one of the pitfalls in previous studies.

The rest of this paper is organized as the following: section 2 provides a relevant theoretical framework. In section 3, we introduce the structural model. Section 4 discusses empirical strategy and econometric issues. Section 5 presents the data and descriptive statistics. The results and discussion are in section 6, and the last section concludes the paper.

2. Theoretical framework

Productivity – Export

Not all firms engage in export activities. Eaton, Kortum, and Kramarz (2004) find that among French manufacturing firms, exporters are a minority that tends to be more productive and larger than non-exporters. The main export barriers are the entry costs. These costs can be variable, with a standard “iceberg” assumption (Bernard, Eaton, Jensen, & Kortum, 2003), or fixed (Melitz, 2003; Roberts & Tybout, 1997).

The variable costs consist mainly of the transportation and tariffs, in which they vary with the amount of export shipment. The fixed costs are the initial costs each firm invests to obtain a permit, establish the distribution network and various other transaction costs.

Among all firms, only those at the upper end of the productivity distribution can afford these costs of entering the foreign markets. But the initial productivity of each firm and the productivity distribution is exogenously determined (Melitz, 2003) or depend only on the variation of the efficiency of firm (Bernard et al., 2003). The theories developed thereafter have largely neglected it and thus the source of productivity gain remains a “black box”.

There are attempts to formally model the firm export with the endogenous productivity. Ederington and McCalman (2008) develop a dynamic model with endogenous firm-level productivity by using an adoption of new technology to explain heterogeneity in firm's productivity. In this model, the difference across firms is the timing of adoption due to the high cost, albeit marginally decreasing, of early technology adoption.

Segerstrom and Stepanok (2011) propose a quality ladders endogenous growth model without Melitz-type assumptions. Instead, they distinguish two types of R&D technologies: inventing higher quality of existing products, and learning how to export. The latter involves an investment in terms of a stochastic fixed market entry cost. Compared to Melitz (2003), the productivity threshold does not exist in this setting and there is an overlap between exporters' and non-exporters' productivity distribution. The difference between this quality ladders model and the model by Ederington and McCalman (2008) is that each product requires different level of R&D. It is therefore more difficult to invest in R&D and learn how to export complex and highly advanced products. Restated, the difference in product quality versus the difference in timing of technology adoption.

Innovation – Productivity

Although the study of the innovation effect on productivity is mostly empirical (see Cohen, 1995; Griliches, 2000 for some surveys), it is the endogenous growth theory that provides a theoretical foundation linking economic output and innovation (P. A. Aghion, Howitt, Brant-Collett, & Peñalosa, 1998; Howitt, 2000; Romer, 1990). Citing Romer (1990), *“technological change provides the incentive for continued capital accumulation, and together ... account for much of the increase in output per hour worked.”*

On a firm level, Klette and Kortum (2004) develop further the theoretical model to link firm's heterogeneity, R&D, and productivity based on several stylized facts from empirical studies on the subject. In this model, the heterogeneity of productivity is derived from a variation in the size of innovation steps. The implication of the model then predicts that R&D intensity is positively correlated with persistent differences in productivity across firms.

3. Structural model

The seminal work of Crépon et al. (1998) (for short, CDM) provides a structural model framework in order to analyze the relationship between R&D, innovation, and productivity at a firm level. The setup is similar to equations 1-4 below. In addition, we include firm's export as the last equation. The structure can be formulated as:

$$g^* = \beta_{0,1} + \sum_n \beta_{n,1} x_{n,1} + \varepsilon_1 \quad (1)$$

$$k^* = \beta_{0,2} + \sum_m \beta_{m,2} x_{m,2} + \varepsilon_2 \quad (2)$$

$$i = \beta_{0,3} + \beta_k k + \beta_{IMR} IMR + \sum_l \beta_{l,3} x_{l,3} + \varepsilon_3 \quad (3)$$

$$p = \beta_{0,4} + \beta_i i + \beta_e e + \beta_{IMR} IMR + \sum_j \beta_{j,4} x_{j,4} + \varepsilon_4 \quad (4)$$

$$e = \beta_{0,5} + \beta_p p + \sum_s \beta_{s,5} x_{s,5} + \varepsilon_5 \quad (5)$$

where g^* denotes innovation input propensity (a dummy with value 1 if total innovation investment is positive⁴), k^* denotes innovation input intensity (logged total innovation investment per employee), i denotes innovation output (logged innovative sales per employee), p denotes productivity (logged value added per employee), and e denotes export intensity (logged export value per employee), β 's are parameters to estimate, and ε is the disturbance term.

Equation (1) examines the decision of the firm to invest in innovation input. Here, $x_{n,1}$ is a vector of the independent variables explaining the decision of the firm to invest in innovation. These variables include physical capital (logged sum of building, machinery, and inventories), human capital (fraction of higher educated employees), and ownership structure variables (categorical variables indicating a firm as being non-affiliated, part of a uninational corporate group, domestic MNEs, or foreign MNEs)⁵. The non-affiliated firms are the reference group.

Equation (2) considers the amount of total innovation investment. Again, $x_{m,1}$ is a vector of explanatory variables, which is the same as $x_{n,1}$ with firm size (measured as logged number of employees) as an additional variable⁶. Equations (1) and (2) are called *innovation input equations*.

Then, equation (3) explains the innovation output of the firm, or the *innovation production function*. The predicted value of innovation input from the previous equation (k) is used as one of the regressors. A vector of explanatory variables, $x_{l,3}$, determines innovation output and includes similar variables as $x_{m,1}$ with an addition of Cooperation variable, indicating if a firm has formal cooperation agreements with external parties or not. The *IMR* variable is the inverted Mills ratio, used to correct for selection bias (Heckman, 1979).

Equation (4) explains firm's productivity. The predicted value of innovation output from the third equation (i) is used as a regressor. Similarly, $x_{j,4}$ include the same variables as $x_{n,1}$. Export intensity (e) is used as a simultaneous explanatory variable to allow for an interaction between productivity and export. This way we test both the self-selection and learning-by-exporting hypotheses. The self-selection expects firm's productivity to later export, while learning-by-exporting expects the reverse – firms learn from previous export to become more productive in later periods (Andersson & Lööf, 2009; Clerides, Lach, & Tybout, 1998). Eventually, equation (5) explains the export performance. The $x_{s,5}$ vector of explanatory variables include, in addition to $x_{j,4}$, a dummy indicating previous export experience. Productivity (p) is used as an additional regressor, to interact with equation (4).

⁴ Total innovation investment is the sum of six innovation investment categories: engagement in intramural R&D, engagement in extramural R&D, engagement in acquisition of machinery, engagement in other external knowledge, engagement in training, and engagement in market introduction of innovation (OECD, 2005).

⁵ This categorical variable for ownership structure is a registered data obtained from Statistics Sweden. We prefer using this categorical variable to a dichotomous variable in CIS data indicating whether a firm belongs to a group or not. This type of substitution is argued to be useful for improving the quality of an empirical analysis in CIS data (Mairesse & Mohnen, 2010).

⁶ Unfortunately, the data does not allow us to construct the market share variable, which is common in the study of this kind (see how to construct the variable in Crépon et al. (1998)). It is expected to be positively related to R&D.

4. Empirical strategy

In this paper, we employ two alternative estimation strategies in order to test the relationship between innovation input, innovation output, productivity, and export (the system of equations in section 2). The first and preferred strategy is a three-step procedure. The second alternative is a two-step procedure in which we present the results in the Appendix. The three-step procedure is as follows:

Step 1: Innovation Input's determinants (Generalized Tobit model)

In this step, we estimate innovation input equations, equations (1) and (2), simultaneously by the Generalized (Type2) Tobit, or sometimes called Heckman selection model (Heckman, 1979). This implies that equation (1) is the selection equation in which the innovation input propensity is the dependent variable. The outcome equation is, therefore, equation (2) and the innovation input intensity is the dependent variable. We use the Full-Information Maximum Likelihood estimator (FIML) to jointly estimate the two equations. Alternatively, we could estimate them as a two-step procedure, with the first being probit, and the second being Ordinary Least Square (OLS) with the inclusion of the Inverted Mills ratio. However, following Verbeek (2008), the first approach is superior to the two-step approach in terms of consistency and efficiency. The general practice for consistent estimates is that at least one of the independent variables should be dropped from the outcome equation. For this paper, we decide to drop firm size. This follows the stylized fact that R&D intensity is independent of firm size (Klette & Kortum, 2004).

Step 2: Innovation Input to Innovation Output (OLS)

In this step, we estimate innovation production function, equation (3), using Ordinary Least Square (OLS). The choice of estimator is in line with recent advancement in estimating the knowledge production function, which does not assume interaction between innovation output and productivity, leaving OLS as a safe estimator (Mairesse & Robin, 2012). From the first step, we use the lagged predicted value of the dependent variable, namely the innovation input intensity (of year 2004), as one of the main independent variable in this step to explain innovation output (in 2006). There are three reasons for this: (i) in order to link step 1 with step 2 as part of system of equations (ii) in order to reduce the potential endogeneity problem by replacing the observed value of innovation input with its predicted value as an instrument⁷ (iii) in order to reduce reverse causality and endogeneity problems by using the two years lag of the main independent variable (innovation input in 2004) to explain dependent variable (innovation

⁷ The main explanatory variable for innovation output is innovation input (intensity). However, this variable is argued to be potentially an endogenous variable, since unobserved characteristics could increase both firms' innovation input efforts and its innovativeness (Mairesse & Robin, 2012; Mohnen, Mairesse, & Dagenais, 2006). Furthermore, it is suggested that predicted value (instead of observed value) can act as an instrument and reduce endogeneity problem (Lööf & Heshmati, 2006; Mairesse & Robin, 2012).

output in 2006)⁸. Standard errors are bootstrapped to correct for the bias induced by an inclusion of predicted regressor (Antonietti & Cainelli, 2011; Mairesse & Robin, 2012).

An important issue is that this step limits the observations to a subsample of innovative firms (378 firms), since we are seeking to explain the innovation output “of innovative firms.” Innovative firms are defined as the firms which have positive innovation input (total investment in innovation activities) and positive innovation output (innovative sales) (Lööf & Heshmati, 2006). However, this may cause selectivity bias, since the total sample is reduced to the non-random subsample of innovative firms (Crépon et al., 1998; Lööf & Heshmati, 2006). In order to deal with the selectivity bias, we use the Inversed Mills ratio variable calculated from the first step to be included as an additional regressor in this and the next step.

Step 3: Innovation Output to Productivity & Export (3SLS)

In the third and final step, we simultaneously estimate productivity and export equations, equations (4) and (5), using the Three-Stage Least Square estimator (3SLS)⁹. There are two reasons for the chosen estimator: (i) in order to deal with simultaneity problem, and (ii) in order to allow for an interaction between productivity and export to test for the two non-mutually exclusive hypotheses concerning productivity and export links, namely self-selection versus learning-by-exporting. An alternative estimator is Two-Stage Least Square (2SLS), but 3SLS has higher efficiency advantage over 2SLS by taking into account the correlations of the error terms between equations (Greene, 2003). We use the lagged predicted value of the innovation output from the previous step as an independent variable (in 2006) to explain subsequent productivity. The logic is the same as the three reasons for including predicted innovation input to explain innovation output, provided in step 2.

An alternative strategy for step 3 is to allow an interaction between innovation output, productivity, and export altogether and reduce the estimation down to two steps (the result is also reported in the Appendix). We employ the three-step approach for the main findings in this paper. The argument is that, theoretically, it is not widely accepted to assume that innovation and productivity can influence each other in both directions. Only an innovation-productivity direction is theoretically and empirically established (Crépon et al., 1998; Lööf & Heshmati, 2006; Mairesse & Mohnen, 2010). For the innovation-export, although the incentive to invest in R&D from being an exporter is formulated theoretically (Long et al., 2011), the evidences are still not abundant. As for econometric issues, although both the three- and two-step procedures take care of the selectivity and simultaneity biases, we also prefer the three-step approach. Because the three-step procedure has an additional dynamic structure that allows for a proper lag – not only from innovation input to innovation output, as in the two-step procedure, but also from

⁸ There are good reasons to believe that innovation input (investments) takes time to exhibit its effect on innovation output (innovative sales). Due to the limitation of data, this lag structure is seldom considered in CIS data. However, since we merge the two waves of CIS, this paper is able to use the lag in the analysis.

⁹ Here, productivity and export are endogenous variables and the other variables are exogenous in this step.

innovation output to productivity and export. In this way, the problem of endogeneity and reverse causality is substantially reduced.

More specifically, the main difference between the three- versus two-step procedure is that in the three-step procedure, the predicted values from the steps 1 & 2 are included as independent variables in the succeeding steps, 2 & 3, to reduce the simultaneity bias. The lagged structure of the independent variables takes into account the endogeneity issue. The productivity-export equations allow for an interaction between productivity and export. This can help us determine the relationship direction (self-selection vs. learning-by-exporting).

5. Data

Our main data for the analysis comes from the Community Innovation Survey (CIS) for Sweden. It contains two waves: CIS4 survey for the years 2002-2004, CIS2006 survey for the years 2004-2006. The surveys are conducted by Statistics Sweden.

We complement the CIS data by including the annual firm's registry and export dataset by matching the encoded unique firm identification number. Therefore, we are able to construct a panel dataset for firms that appear in both CIS waves to have a range from 2002-2008. The main advantage for this setup is that we are able to analyse the structural model in a dynamic setting.

We restrict our analysis to the manufacturing sector to focus on firms that export what they actually produce. The sample includes 1,718 firms in total. The descriptive statistics for all variables in this paper are presented in table 1 below. Table 6 in the Appendix lists the correlation of all variables and multicollinearity seems not to be a problem.

Table 1: Descriptive statistics

VARIABLES	Obs.	Mean	Std. Dev.	Minimum	Maximum
Physical capital	11,569	171,449,800	11,524,990	742	226,898,500
Human capital	11,729	0.135	0.179	0	1
Firm size	11,725	219.930	1,195.400	1	39,554
Uninational	12,026	0.315*	0.465	0	1
Domestic MNEs	12,026	0.217*	0.412	0	1
Foreign MNEs	12,026	0.219*	0.413	0	1
Cooperation	1,917	0.497*	0.500	0	1
Innovation Input (Propensity)	12,026	0.127*	0.333	0	1
Innovation Input	3,426	43,768	474,383	0.045	22,580,650
Innovation Output	659	1,882,750	3,587,762	3,101	92,110,160
Productivity	11,489	754,182	2,996,460	2,546.56	278,207,200
Export Intensity	3,355	751,735	1556093	0.008	28,223,150
Export Experience	6,801	0.767*	0.423	0	1

* The fraction of observations with value 1.

6. Results and discussion

As we elaborate in section 3, the structural model is estimated in three steps: (1) the innovation input equation (Heckman), (2) the innovation production function (OLS), and (3) the productivity & export performance equation (3SLS). We present the results of steps 1 to 3 in tables 2 to 4, respectively.

In table 1, the joint estimation of innovation input equations, i.e. equations (1) and (2), is reported. Column (1) is the selection equation corresponding to equation (1). The dependent variable is innovation input propensity, measured as a dummy with value one if total investment in innovation activities are positive from 2002 to 2004 (denoted by *2004*) and from 2004 to 2006 (denoted by *2006*). Column (2) is the outcome equation corresponding to equation (2). The dependent variable is innovation input intensity, measured as logged total investment in innovation activities per employee, observed in the same time as innovation input propensity.

Table 2: Step 1 - Innovation Input determinants

Dependent variable: Innovation Input (*logged Total innovation investment per employee*)

Estimation: Generalized Tobit model (Heckman using FIML)

VARIABLES	(1) Selection Equation Innovation Input <i>2004, 2006</i> (Propensity)	(2) Outcome Equation Innovation Input <i>2004, 2006</i> (Intensity)
Firm Size <i>(log) (2004, 2006)</i>	0.235*** (0.030)	
Physical Capital <i>(log) (2004, 2006)</i>	0.122*** (0.024)	0.330*** (0.040)
Human Capital <i>(log) (2004, 2006)</i>	1.288*** (0.314)	3.755*** (0.431)
Uninational	0.114 (0.079)	0.180 (0.141)
Domestic MNEs	0.277*** (0.097)	0.219* (0.133)
Foreign MNEs	0.077 (0.099)	0.181 (0.137)
Constant	-3.416*** (0.531)	-2.749* (1.513)
Observations	2,135	-
Uncensored Obs.	-	1,244
Industry dummies	YES	YES
Year dummies	YES	YES

Robust standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 2 shows that both physical capital and human capital have positive and significant influence on both the decision and the intensity of innovation input. This is in line with previous studies using CIS data (Crépon et al., 1998; Lööf & Heshmati, 2006). Ownership structure variables are significant only for domestic MNEs, meaning that there is a significant difference if a firm belongs to Swedish MNEs or not when it comes to innovation decisions, while it makes no difference for firms belonging to a unination corporate group and foreign MNEs.

The next table reports the estimation of innovation production function, equation (3), for innovative firms using OLS. The dependent variable is innovation output, measured as logged innovative sales per employee from 2004 to 2006 (denoted by *2006*). The main independent variable is lagged innovation input from 2002 to 2004 (denoted by *2004*), which is predicted from the previous step (from Table 2).

Table 3: Step 2 - Innovation Input to Innovation Output

Dependent variable: Innovation Output (*logged Innovative Sales per employee*)

Estimation: OLS with the predicted value of Innovation Input from step 1

VARIABLES	(3) Innovation Output <i>2006</i>
Innovation Input <i>(Predicted) (lagged: 2004)</i>	0.333* (0.187)
Firm Size <i>(log)(lagged: 2004)</i>	0.370*** (0.112)
Physical Capital <i>(log)(lagged: 2004)</i>	0.347*** (0.104)
Human Capital <i>(log)(lagged: 2004)</i>	3.068*** (0.914)
Cooperation <i>(lagged: 2004)</i>	0.038 (0.152)
Uninational	0.199 (0.249)
Domestic MNEs	0.730*** (0.278)
Foreign MNEs	0.463* (0.238)
Inverted Mills ratio <i>(2006)</i>	5.071*** (1.466)
Constant	0.333* (0.187)
Observations (innovative firms)	378
R-squared	0.131
Industry dummies	YES

Bootstrapped standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 3 shows that past innovation input has positive (but weakly) significant effect on innovation output, as expected. Firm size and capitals also have positive and significant effects on innovation output of innovative firms. Two of the ownership structure variables, domestic and foreign MNEs, are significant and positive.

The last step of the estimation is reported in table 4. Here, we estimate equations (4) and (5) simultaneously. The chosen estimator is 3SLS (as discussed in section 3).

The predicted lagged innovation output in 2006 is used as the main explanatory variable for productivity. This variable is obtained from the second step (table 3). The dependent variables are productivity and export performance measured as total export value during 2008-2009, columns (4) and (5).

The positive and highly significant result for productivity suggests that it can explain the export performance, which is in line with self-selection literature. On the other hand, export seems to be unable to explain productivity, which means this paper does not find evidence to support learning-by-exporting argument. For past export experience, it is positive and highly significant in explaining the current export performance. Furthermore, physical capital and human capital are positive and significant to explain productivity, as expected. Ownership structure variables, i.e. domestic and foreign MNEs, turn out to be positive and significant to explain only export performance.

Moving backward from step 3 to 1, tables 4 to 2, the system of equations adopted in this paper reveals the mechanics behind the export behavior of firms. Productive firms become exporters and there is no evidence that export leads a firm to be productive. The source of such productivity heterogeneity is innovative performance (output) of firms in the past. Lastly, such innovative output is the result of the amount of investment in innovation activities and the decision to invest in the past.

As an alternative approach, we allow for an interaction between innovation output, productivity, and export in a simultaneous framework. Table 5 in the Appendix lists the results. We first estimate the three equations, eqs. (3) – (5), separately by OLS, then simultaneously by 3SLS. The Hausman test rejects the null hypothesis of simultaneity. The results suggest that lagged innovation input can explain current innovation output. Productivity can explain export intensity. Interestingly, export intensity is positive and significant in explaining both innovation output and productivity.

Table 4: Step 3 - Innovation Output to Productivity & Export**Dependent variables:** Productivity: (*logged Value Added per employee*), Export: (*logged Export value per employee*)**Estimation:** Three-stage Least Square (3SLS)

VARIABLES	(4) Productivity (2008-2009)	(5) Export Intensity (2008-2009)
Innovation Output (<i>Predicted</i>)(<i>log</i>) (2006)	0.161** (0.074)	
Productivity (<i>log</i>) (2008-2009)		2.017*** (0.770)
Export Intensity (<i>log</i>) (2008-2009)	-0.027 (0.080)	
Export Experience (2006)		2.867** (1.135)
Size (<i>log</i>) (2008-2009)	0.017 (0.032)	0.025 (0.099)
Physical Capital (<i>log</i>) (2008-2009)	0.084*** (0.031)	
Human Capital (<i>log</i>) (2008-2009)	0.614* (0.348)	
Uninational	0.085 (0.083)	0.070 (0.484)
Domestic MNEs	0.135 (0.155)	1.400*** (0.492)
Foreign MNEs	0.169 (0.174)	1.532*** (0.554)
Inverted Mills ratio (2006)	-0.071 (0.178)	
Constant	10.510*** (0.913)	-17.994* (9.485)
Observations	435	435
R-squared	0.102	0.174
Industry dummies	YES	YES

 Bootstrapped standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

7. Conclusion

It is well known that exporters are productive firms. But the source of their productivity is left unexplained. This paper aims to endogenize the productivity heterogeneity of exporting firms by incorporating innovation in a structural model framework. In doing so, we close the gap between the innovation-productivity and productivity-export literature.

There are two novelties in this paper. Firstly, we open the black box concerning the source of productivity heterogeneity of exporting firms. Although we know that productive firms can become exporter later, we answer to the question why those productive firms are productive in the first place. Secondly, by merging the two waves of Swedish CIS data and tracing the participants' behavior from 2002 to 2009, it becomes possible to (i) consider lagged value of innovation input to explain future innovation output, and then (ii) consider the innovation output to explain further future productivity and export performance. It implies that we (i) give enough time for innovation input to impact on innovation output and also (ii) give enough time for innovation output to have its effect on productivity and export performance. Such a dynamic structure can substantially reduce the reverse causality and endogeneity bias.

The main findings are that exporters are productive firms, which cross the productivity threshold. They have passed it because they succeed in appropriating the innovation output in the past, which is driven in turn by the decision and amount of investment in various investments in innovation activities.

Appendix

Table 5: Step 2 - Innovation Output, Productivity, and Export Performance

Dependent variables: *logged Innovative sales per employee*, *logged Value added per employee*, and *logged Total Export value per Employee*

Estimation: Equation-by-equation OLS, Simultaneous equations (3SLS) with bootstrapped standard errors.

VARIABLES	(6)	(7)	(8)	(9)	(10)	(11)
	Innovation Output	OLS (2006) Productivity	Export Intensity	Innovation Output	3SLS (2006) Productivity	Export Intensity
Innovation Input <i>(Predicted) (logged: 2004)</i>	0.292** (0.129)			0.288* (0.151)		
Innovation Output <i>(log) (2006)</i>		0.011 (0.056)			0.007 (0.062)	
Productivity <i>(log) (2006)</i>			4.388*** (0.813)			4.414*** (0.764)
Export Intensity <i>(log) (2006)</i>	0.197* (0.104)	0.122*** (0.023)		0.199* (0.112)	0.123*** (0.029)	
Export Experience <i>(2006)</i>			2.393 (2.574)			2.370 (2.496)
Process Innovation <i>(2006)</i>		0.005 (0.036)			0.004 (0.042)	
Size <i>(log) (2006)</i>	0.016 (0.057)	0.012 (0.016)	-0.175*** (0.063)	0.015 (0.062)	0.012 (0.012)	-0.176*** (0.066)
Physical Capital <i>(log) (2006)</i>	0.021 (0.074)	0.059*** (0.020)		0.019 (0.067)	0.059*** (0.018)	
Human Capital <i>(2006)</i>	1.335* (0.744)	0.617*** (0.214)		1.318* (0.758)	0.620** (0.261)	
Cooperation <i>(2006)</i>	0.098 (0.114)			0.096 (0.142)		
Uninational	-0.207 (0.204)	-0.005 (0.079)	0.036 (0.472)	-0.209 (0.255)	-0.006 (0.078)	0.035 (0.401)
Domestic MNEs	-0.269 (0.238)	-0.027 (0.080)	1.110*** (0.418)	-0.277 (0.296)	-0.029 (0.105)	1.105** (0.430)
Foreign MNEs	-0.393 (0.262)	-0.053 (0.082)	1.230*** (0.406)	-0.399 (0.285)	-0.056 (0.103)	1.225*** (0.438)
Inverted Mills ratio <i>(2006)</i>	1.360* (0.777)			1.322 (0.873)		
Constant	7.161*** (1.694)	10.890*** (0.730)	-48.264*** (11.116)	7.210*** (2.152)	10.924*** (0.775)	-48.584*** (10.292)
Observations	328	328	328	328	328	328
R-squared	0.131	0.328	0.406	0.131	0.328	0.406
Industry dummies	NO	NO	NO	NO	NO	NO

Bootstrapped standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1
*Note: Hausman test of OLS vs. 3SLS reject the null hypothesis of simultaneity.

Table 6: Correlation table

	Physical Capital	Human Capital	Firm Size	Produc- tivity	Uni national	Domestic MNEs	Foreign MNEs	Coope- ration	Inno. Input Propensity	Inno. Input	Inno. Output	Export Intensity	Export Experience
Physical Capital	1.000												
Human Capital	-0.201***	1.000											
Firm Size	0.255***	0.093*	1.000										
Productivity	0.292***	0.231***	0.276***	1.000									
Uninational	-0.047	-0.099*	-0.302***	-0.163***	1.000								
Domestic MNEs	-0.004	0.131**	0.139**	0.056	-0.394***	1.000							
Foreign MNEs	0.075	0.037	0.343***	0.211***	-0.382***	-0.480***	1.000						
Cooperation	0.147***	0.210***	0.292***	0.234***	-0.096*	0.119**	0.025	1.000					
Inno. Input Propensity	0.122**	0.014	0.062	0.022	0.036	0.041	-0.047	0.173***	1.000				
Inno Input	0.123**	0.336***	0.016	0.232***	-0.037	0.103*	-0.072	0.281***	0.507***	1.000			
Inno. Output	0.060	0.162***	0.141**	0.297***	-0.169***	0.058	0.121**	0.170***	-0.002	0.228***	1.000		
Export Intensity	0.198***	-0.075	0.099*	0.239***	-0.268***	0.214***	0.105*	0.166***	0.114*	0.257***	0.253***	1.000	
Export Experience	0.031	0.031	0.134**	0.102*	-0.201***	0.119**	0.077	0.080	-0.006	0.092*	0.053	0.296***	1.000

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

References

- Aghion, P., & Howitt, P. (1992). A Model of Growth Through Creative Destruction. *Econometrica*, 60(2), 323-351. doi: 10.2307/2951599
- Aghion, P. A., Howitt, P. A., Brant-Collett, M., & Peñalosa, C. G. (1998). *Endogenous Growth Theory*. MIT Press.
- Andersson, M., & Lööf, H. (2009). Learning-by-Exporting Revisited: The Role of Intensity and Persistence*. *Scandinavian Journal of Economics*, 111(4), 893-916. doi: 10.1111/j.1467-9442.2009.01585.x
- Antonietti, R., & Cainelli, G. (2011). The role of spatial agglomeration in a structural model of innovation, productivity and export. *Annals of Regional Science*, 46, 577-600.
- Bernard, A. B., Eaton, J., Jensen, J. B., & Kortum, S. (2003). Plants and Productivity in International Trade. *The American Economic Review*, 93(4).
- Bernard, A. B., & Jensen, J. B. (1999). Exceptional exporter performance: cause, effect, or both? *Journal of International Economics*, 47(1), 1-25.
- Bernard, A. B., & Wagner, J. (1997). Exports and Success in German Manufacturing. *Review of World Economics*, 133(1), 134-157.
- Clerides, S. K., Lach, S., & Tybout, J. R. (1998). Is Learning by Exporting Important? Micro-Dynamic Evidence from Colombia, Mexico, and Morocco. *The Quarterly Journal of Economics*, 113(3), 903-947. doi: 10.1162/003355398555784
- Cohen, W. M. (Ed.). (1995). *Empirical Studies of Innovative Activity*. Oxford: Blackwell.
- Crépon, B., Duguet, E., & Mairesse, J. (1998). Research investment, innovation and productivity: An econometric analysis. *Economics of Innovation and New Technology*, 7(2), 115-158.
- Delgado, M. A., Fariñas, J. C., & Ruano, S. (2002). Firm Productivity and Export Markets: A Non-Parametric Approach. *Journal of International Economics*, 57(2), 397-422.
- Eaton, J., Kortum, S., & Kramarz, F. (2004). Dissecting trade firms, industries, and export destinations *Federal Reserve Bank of Minneapolis, Research Department staff report 332* Retrieved from <http://woodrow.mpls.frb.fed.us/research/sr/sr332.html>
- Ederington, J., & McCalman, P. (2008). Endogenous firm heterogeneity and the dynamics of trade liberalization. *Journal of International Economics*, 74(2), 422-440. doi: <http://dx.doi.org/10.1016/j.jinteco.2007.07.001>
- Greene, W. H. (2003). *Econometric analysis*: Prentice Hall.
- Griliches, Z. (2000). *R&D, Education, and Productivity: A Retrospective*. Cambridge, Mass.: Harvard Univ. Press.
- Hall, B. H., & Mairesse, J. (2006). Empirical studies of innovation in the knowledge-driven economy. *Economics of Innovation and New Technology*, 15(4/5), 289-299.
- Heckman, J. J. (1979). Sample Selection Bias as a Specification Error. *Econometrica*, 47(1), 153-161.
- Howitt, P. (2000). Endogenous Growth and Cross-Country Income Differences. *The American Economic Review*, 90(4), 829-846.
- Klette, T. J., & Kortum, S. (2004). Innovating Firms and Aggregate Innovation. *Journal of Political Economy*, 112(5), 986-1018.

- Long, N. V., Raff, H., & Stähler, F. (2011). Innovation and Trade with Heterogeneous Firms. *Journal of International Economics*, 84(2), 149-159.
- Lööf, H., & Heshmati, A. (2006). On the relationship between innovation and performance: A sensitivity analysis. *Economics of Innovation and New Technology*, 15(4/5), 317-344.
- Mairesse, J., & Mohnen, P. (Eds.). (2010). *Using innovation surveys for econometric analysis* (Vol. 2). London: Burlington Academic Press.
- Mairesse, J., & Robin, S. (Eds.). (2012). *The importance of Process and Product Innovation for Productivity in French Manufacturing and Service Industries*. Oxford: Oxford University Press.
- Melitz, M. J. (2003). The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity. *Econometrica*, 71(6), 1695-1725.
- Mohnen, P., Mairesse, J., & Dagenais, M. (2006). Innovativity; A comparison across seven European countries. *Economics of Innovation and New Technology*, 15(4/5), 391-413.
- OECD. (2005). Oslo Manual: guidelines for collecting and interpreting innovation data. Paris: OECD and Eurostat.
- Roberts, M. J., & Tybout, J. R. (1997). The Decision to Export in Colombia: An Empirical Model of Entry with Sunk Costs. *The American Economic Review*, 87(4), 545-564.
- Romer, P. (1990). Endogenous Technological Change. *Journal of Political Economy*, 98, 71-102.
- Segerstrom, P., & Stepanok, I. (2011). Learning How to Export *Working Paper Series in Economics and Finance*. Stockholm School of Economics.
- Wagner, J. (2007). Exports and Productivity: A Survey of the Evidence from Firm-level Data. *World Economy*, 30(1), 60-82. doi: 10.1111/j.1467-9701.2007.00872.x
- Verbeek, M. (2008). *A guide to modern econometrics* (3 ed.). West Sussex: John Wiley & Sons.