Product and Process Innovation and the decision to Export: Firm-level evidence for Belgium

I. Van Beveren and H. Vandenbussche

Discussion Paper 2009-32

Institut de Recherches Économiques et Sociales
de l'Université catholique de Louvain

UCL

Product and Process Innovation and the decision to Export: Firm-level evidence for Belgium

Ilke Van Beveren*†** and Hylke Vandenbussche†‡§

This version: September 14, 2009

Abstract**

Using data from the Community Innovation Survey for Belgium in two consecutive periods, this paper explores the relationship between firm-level innovation activities and the propensity to start exporting. To measure innovation, we include indicators of both innovative effort (R&D activities) as well as innovative output (product and process innovation). Our results suggest that the combination of product and process innovation, rather than either of the two in isolation, increases a firm's probability to enter the export market. After controlling for potential endogeneity of the innovation activities, only firms with a sufficiently high probability to start exporting engage in product and process innovation prior to their entry on the export market, pointing to the importance of self-selection into innovation.

Key words: Exports, Product innovation, Process innovation, Self-selection, Firm heterogeneity.

JEL Classification: D24, F14, L25, O31, O33

* Lessius University College, Department of Business Studies, Antwerp, Belgium.
† Université Catholique de Louvain, Belgium.
‡ Katholieke Universiteit Leuven, LICOS Centre for Institutions and Economic Performance, Belgium.
§ CEPR.
** **Acknowledgements**
We are thankful to Bee Aw, Mark Roberts, Bruno Cassiman, Stijn Vanormelingen, Davide Castellani, Maurizio Zanardi and Marton Csillag for helpful discussions and suggestions. Participants of the International Economics Group Seminar Series at UCL, the International Trade Workshop in Brussels and the European Trade Study Group conference in Rome provided useful comments. We would also like to thank Manu Monard, Peter Teirlinck and the CFS-STAT Commission for allowing us to access the CIS data for Belgium, for answering questions related to the data and for their hospitality during visits there. While we are thankful to all of them, none of them are responsible for any errors we have made.

1. Introduction

There is a large and growing body of literature dealing with the link between firms' decision to export and their productivity. The seminal works of Bernard and Jensen (1999) and Melitz (2003) have shown that only the more productive firms self-select into exporting, since only firms with an efficiency level above a certain threshold, are able to overcome the fixed costs associated with entry on the export market. The theoretical self-selection literature typically assumes that firms' productivity is a random, exogenous draw from a Pareto distribution.

More recent contributions to the literature (e.g. Yeaple, 2005; Bustos, 2005) have sought to endogenize firm-level productivity, hence allowing for the possibility that firms can influence their own efficiency level, rather than simply observing it in each consecutive period. One of the ways in which firms can increase their productivity, is through innovation activities. In the theoretical framework of Yeaple (2005), firms have the possibility to adopt either a high-technology, low unit cost or low-technology, high unit cost production process. The low-unit cost technology entails a higher fixed cost of technology adoption. In the presence of fixed costs to enter the export market, only those firms that adopt the low unit-cost technology will be able to start exporting.

In response to these developments, several authors have explored the relationship between firms' innovation activities and their propensity to engage in exports. However, thus far, the empirical results on the link between innovation activities and the firm's export decision have been mixed. Moreover, results seem to depend on which innovation measures are used. Specifically, both Aw, Roberts and Winston (2007), as well as Cassiman and Martinez-Ros (2007) fail to find a significant link between firm-level R&D (innovative effort) and the probability of firms to start exporting, using firm-level data on manufacturing firms in Taiwan and Spain respectively.

When innovation output measures are considered, the link between innovation and firms' propensity to export appears to yield stronger, but mixed results. While Caldera (2009) and Cassiman and Martinez-Ros (2007) both use the Spanish ESEE data set, their analysis yields different findings. In particular, Cassiman and Martinez-Ros (2007) identify product innovation, but not process innovation, as a driver of firm-level export propensity; while Caldera (2009) finds both product and process innovation to matter, although the impact of product innovation is higher than that of process innovation. These results are relatively robust to several endogeneity controls and other robustness checks. Damijan, Kostevc and Polanec (2008) on the other hand, using data on the Slovenian manufacturing sector and applying matching techniques to account for the endogeneity of the innovation activities, find no evidence that product or process innovation acts as a significant driver of export propensity at the firm level. They do provide evidence that firms engage significantly more in process innovation after entering the export market. Finally, Becker and Egger (2007) apply matching techniques to German survey data find that firms introducing a product *and* process innovation simultaneously increase their propensity to export by about ten percentage points.

Product innovation acts as a significant driver of firms' export propensity when introduced in isolation, but process innovation does not.

The present paper aims to explore this link between innovation and firms' export propensity further using data from the Community Innovation Survey (CIS) for Belgium. The CIS questionnaire is carried out in all European countries every four years. In the empirical analysis, we will use CIS data for two consecutive periods, CIS3 (1998-2000) and CIS4 (2002-2004[1]). The data contain detailed information on firms' innovation characteristics, and also record a number of other firm-level variables, such as firms' export status and intensity. Firm-level data can only be obtained through the national statistical offices. The richness of the innovation measures available in the CIS data and the lack of other firm-level panel data sets containing detailed innovation characteristics[2] has resulted in a growing number of empirical papers that use the CIS data in recent years, examples include Castellani and Zanfei (2006, 2007), Damijan et al. (2008), Griffith, Mairesse and Peters (2006) and Mairesse (2004).

Several endogeneity issues need to be addressed when analyzing the link between firms' innovation activities and their propensity to export. First, since firms typically make their innovation and export decisions simultaneously, a *simultaneity bias* emerges. Second, since exporting activities tend to exhibit persistence over time (Aw et al., 2007), a *causality bias* arises when past exporting history is not properly controlled for. Finally, to the extent that firms can anticipate entry into the export market and their innovative efforts are driven by this future prospect, the introduction of new innovations is endogenous to firms' export decision (Costantini and Melitz[3], 2007). This *anticipation effect* is the third source of endogeneity when analyzing the link between firm-level innovation and exporting activities.

To account for the simultaneity of firm-level innovation and exporting decisions and to rule out past exporting history, we limit the sample used in the empirical analysis in two ways. First, we focus on firms that have answered the CIS questionnaire in two consecutive periods. This will allow us to use lagged (initial) innovation and other firm-level characteristics as potential determinants of firms' propensity to export. This limits the sample to 600 firms. Second, to control for the persistence of firm-level exports (see for instance Aw et al., 2007), we focus our attention on firms that started exporting in 2004 *(Starters)* and compare these to

[1] The CIS data always cover a period of two years. Firms are asked to report innovation activities between the beginning and end of this period. All financial and accounting information, such as the value of sales, export intensity and total amount spent on R&D pertains only to the last year in the data (2000 and 2004 for CIS3 and CIS4 respectively).
[2] Exceptions include the ESEE data for Spain, which contains information on firm-level product and process innovation and the SPRU database for the UK, which contains detailed information on firm-level innovations since the second World War.
[3] Costantini and Melitz (2007) analyze the joint entry, exit, export and innovation decisions of firms confronted with trade liberalization in a dynamic setting. They find that the anticipation of upcoming liberalization can induce firms to innovate prior to their entry on the export market.

a control group of firms that did not export in either period *(Non-exporters)*. The final sample of firms included in the empirical analysis amounts to 189 firms.

Similar to previous research, we fail to find a link between firm-level internal (or external) R&D and the exporting decision. Moreover, when we add product and process innovations simultaneously as determinants of the firms' exporting decision, our findings are similar to those obtained by Cassiman and Martinez-Ros (2007), i.e. product innovation, but not process innovation, acts as a significant driver of firms' entry on the export market. However, inspection of the data reveals that more than fifty percent of innovating firms in our sample introduce a product and process innovation *simultaneously*, rather than one of the two in isolation.

Specifically, 48 percent of all firms that introduced a product innovation between 1998 and 2000 also introduced a process innovation in our sample. Similarly, 58 percent of all firms that have introduced a process innovation during the same period, simultaneously introduced a product innovation. This results in a correlation between the two variables of 0.4428. When we account for this correlation in the empirical analysis, our results suggest that it is not so much product or process innovation in isolation, but rather the combination of the two, which drives firms into the export market.

Finally, to account for the anticipation effect, we follow Cassiman and Martinez-Ros (2007) and Caldera (2009) and implement an instrumental variable approach, using firm-level innovation inputs (internal and external R&D) and training activities as instruments for the innovation decision. After controlling for potential endogeneity of the innovation activities, we find no evidence that firms engaging in product and/or process innovation are more likely to enter the export market. These results suggest that only firms with a sufficiently high probability to start exporting will engage in product and process innovation prior to their entry on the export market, pointing to the importance of self-selection into innovation activities. These results are in line with those obtained by Damijan et al. (2008) who also fail to find a significant link between product and process innovation and firms' entry on the export market.

Our analysis contributes to the existing literature in a number of important ways. First, unlike most of the existing empirical work[4], we take the correlation between product and process innovation explicitly into account in the empirical analysis, hence allowing for potential complementarities between the two innovation types. Furthermore, our results point to the importance of accounting for all potential sources of endogeneity of firm-level innovation in the exporting decision. After accounting for the three types of endogeneity outlined above i.e. *simultaneity bias, causality bias* and *anticipation effect*, our results suggest that firms are only more likely to engage in innovation activities if their prospects to enter the export market in the next period are good, i.e. if their future propensity to export is high.

[4] Becker and Egger (2007) are a notable exception.However their analysis includes both starters on the export market and continuing exporters, i.e. they do not control for past exporting history in their analysis.

The rest of the paper is structured as follows. Section 2 reviews the relevant literature, while section 3 discusses the data and reveals interesting empirical facts. Section 4 introduces the empirical model and section 5 presents the empirical results. In section 6 several robustness checks are applied to the data. The final section concludes.

2. Literature review

While the literature on the relationship between firms' participation on export markets and their productivity abounds[5], until recently, it remained largely silent on the sources of the productivity advantages associated with firms' entry on export markets. Following theoretical models of entry and exit (e.g. Jovanovic, 1982; Hopenhayn, 1992; Melitz, 2003), researchers have long continued to assume that the productivity advantage that enabled firms to start exporting (or start producing) was exogenous in nature, hence not determined by any firm-specific effort.

From this early literature dealing with the relationship between exports and productivity, a dual relationship emerges, whereby firms exogenously self-select into the export market (i.e. their productivity is higher than the minimum efficiency level required to enter export markets) and, once they start exporting, have the potential to further increase their productivity through learning effects.

Recently, however, efforts have been made to endogenize firm heterogeneity, allowing firms to engage in productivity-enhancing activities prior to engaging in international markets. Important theoretical contributions in this field include Bustos (2005) and Yeaple (2005). Unlike earlier models of firm dynamics (e.g. Jovanovic, 1982), in Yeaple's model firms are born identical. After being born, they have the possibility to adopt a high-technology, low unit cost production technology, or a low-technology high unit cost technology. In the presence of fixed costs associated with both technology adoption and exporting, the model shows that only those firms adopting the low unit cost technology are able to start exporting. In related work, Costantini and Melitz (2007) analyze the joint entry, exit export and innovation decisions of firms in response to or in anticipation of trade liberalization. Their findings point to the importance of taking the timing and speed of trade liberalization into account when analyzing firms' export and innovation decisions. In particular, they find that anticipation of upcoming trade liberalization and a slow liberalization process can motivate firms to innovate ahead of export market entry.

From these different strands of the literature, three different hypotheses concerning the link between exporting and productivity emerge (Alvarez and Lopez, 2005). Apart from exogenous self-selection and learning effects, firms have the possibility to engage in

[5] For reviews on this extensive literature, we refer to Greenaway and Kneller (2007) and Wagner (2007).

investments aimed specifically at raising their productivity prior to entry on export markets (*Conscious self-selection*). While empirical studies tend to provide evidence in favor of the exogenous self-selection hypothesis, for learning effects the results tend to be mixed.

More recently, several empirical papers provide evidence on the conscious self-selection hypothesis, investigating the link between firms' export propensity and a number of firm-level investments or decisions: training and R&D (Aw et al., 2007), product and process innovation (Damijan et al., 2008) and physical investment (Alvarez and Lopez, 2005; Iacovone and Smarzynska Javorcik, 2008). A common feature all these papers share is that they investigate to what extent certain (investment) activities of firms increase their propensity to engage in exports. Furthermore, all of the studies cited provide evidence on the complementary nature of these investment activities and firms' export propensity.

When investigating the link between firm-level innovation activities and its propensity to (start) export(ing), two types of innovation measures have been used in the literature. Specifically, either innovation input measures, usually expressed as the ratio of R&D over sales or as a dummy variable indicating whether firms engage in R&D, or innovation output measures, typically expressed as dummy variables representing whether firms have introduced a product or process innovation; are used as measures of firm-level innovation activities. As was already noted in the introduction, the impact of firms' innovation activities on their export propensity are mixed and seem to depend on the type of measures used.

Aw, Roberts and Winston (2007) explore the link between firm-level R&D, training, productivity and exports using data on the Taiwanese electronics sector. Their findings suggest that R&D and exporting are not complementary activities, but they have a complementary effect on firm-level productivity. These results seem to imply that the combination of exporting and R&D increases productivity more than the sum of both conducted in isolation. Cassiman and Martinez-Ros (2007) find similar results for the Spanish manufacturing sector, i.e. firms engaging in R&D investment do not exhibit a significantly higher export propensity.

While research spending of firms can be considered a reasonable proxy of firm-level innovative output in the absence of information on the actual innovations firms have introduced, there are several drawbacks associated with the use of R&D spending, which is essentially an input in the innovation production function[6] as a measure of firm-level innovation. First, not all innovation efforts actually lead to the introduction of product or process innovations, i.e. it is possible that firms' efforts to innovate fail for some reason, in which case using R&D rather than actual innovations leads to an overestimation of firms' innovative activities. Second, it is not unlikely that there is a considerable time lag between firms' investment in R&D and the actual introduction of an innovation to the market, in which case the timing of the R&D and innovation decisions do not match, leading to an

[6] See for instance Mairesse and Mohnen (2002).

overestimation of innovation in some years and an underestimation in later years, when the level of R&D spending is lower and innovative output is higher.

Several authors have taken these drawbacks into account and rely on measures of firm-level innovation output rather than inputs to investigate the link between firm-level innovation and export propensity. Becker and Egger (2007) use German survey data, Caldera (2009) and Cassiman and Martinez-Ros (2007) use the ESEE data set for Spain and Damijan et al. (2008) use CIS data for Slovenia to explore the relationship between firm-level innovative output, measured as the introduction of product and process innovations and firms' propensity to (start) export(ing). While these papers share a common purpose and in the case of Caldera (2009) and Cassiman and Martinez-Ros (2009) also use the same data set, some differences between them, both in terms of sample selection, methodology and empirical results are worth noting here.

As shown by Aw et al. (2007) firms' exporting status is characterized by a high persistence, a finding that is also consistent with the prediction of Melitz's model that entry into the export market leads to the incurrence of a fixed cost, which cannot be recovered. Given these preliminaries, it is not unlikely that firms' initial entry versus its continued presence on the export market have different determinants. Iacovone and Smarzynska Javorcik (2008) document, for a sample of Mexican manufacturing firms, an increase in physical investment prior to the introduction of a domestic variety on the export market, but only for new exporters. For firms with prior export experience, no such increase was recorded. These findings point to the importance of taking firms' prior export experience into account in the empirical analysis.

For this reason, Damijan et al. (2008) focus only on first-time exporters when investigating the impact of firm-level innovation activities on firms' propensity to export. While Caldera (2009) and Cassiman and Martinez-Ros (2007) both use the full sample of exporters (starters and firms with export experience) in their analysis, they both perform a number of robustness checks to account for prior experience in exporting. Specifically, Caldera (2009) estimates a dynamic model as a robustness check and Cassiman and Martinez-Ros (2007) repeat their analysis using only starters on the export market versus a control group of non-exporters. In both cases, the main findings are robust to these alternative specifications. Becker and Egger (2007) on the other hand, focus on the full sample of firms and do not differentiate between first-time exporters and continuing exporters. In the empirical analysis below, we follow Damijan et al. (2008) by focusing only on starters on the export market and a control group of non-exporters.

In terms of the methodologies used, the four papers cited above can be divided in two groups. Caldera (2009) and Cassiman and Martinez-Ros (2007) both use a probit model to investigate the relationship between firm-level innovation and export status. To control for unobserved firm heterogeneity, they add random effects to the baseline specification. Apart from a number of control variables, both papers add lagged innovation status for product and process innovation as independent variables. This allows them to control for the simultaneity of the export and innovation decisions. However, while selection on prior export status (i.e. using only starters on the export market) and the use of lagged firm characteristics avoids the pitfalls of persistence in exports and of a simultaneity bias resulting from the timing of the innovation and export decisions, this does not rule out the existence of feedback effects, rendering firm-level innovation endogenous in the export decision framework.

Specifically, if firms have some prior knowledge of their prospects on the export market, they are likely to make their innovation decisions with this prospect in mind. In other words, to the extent that firms can anticipate their entry on the export market and if their innovation efforts are driven by this expectation, product and process innovation cannot be considered exogenous in the export decision. To take this anticipation effect into account, Caldera (2009) and Cassiman and Martinez-Ros (2007) estimate, in addition to their baseline model, several instrumental variables (IV) regressions. Caldera relies on a linear probability framework and uses firm-level funding for innovation as an instrument, while Cassiman and Martinez-Ros rely on IV probit estimation.

Becker and Egger (2007) and Damijan et al. (2008) take a more direct approach to account for the potential endogeneity of the innovation decision in the firm’s exporting decision, both papers apply matching estimators. Becker and Egger (2007) focus on the causal link going from innovation to exporting; while Damijan et al. (2008) look at the bi-directional causal impact. As was noted in the introduction, Becker and Egger (2007) are among the first to take the correlation between firm-level product and process innovation explicitly into account. In their matching analysis, they distinguish between four types of firms: i) firms that did not introduce a product or process innovation, ii) firms that introduced a product, but not a process innovation, iii) firms that introduced a process, but not a product innovation and iv) firms that introduced both a product and process innovation. However, their analysis includes both continuing exporters and starters, i.e. they do not control for past exporting history in their analysis.

In terms of the empirical results, the existing literature remains inconclusive. Caldera (2009), Cassiman and Martinez-Ros (2007) and Becker and Egger (2007) find that the introduction of a product innovation results in an increase in firms’ export propensity. On the other hand, Damijan et al. (2008) find no significant impact of product innovation on the export propensity of Slovenian firms. For process innovation, the findings of Becker and Egger (2007), Cassiman and Martinez-Ros and Damijan et al. (2008) suggest that process innovation does not increase firms’ export propensity. Caldera (2009) on the other hand reports a positive and significant impact of process innovation on the probability of firms to export, a finding that is robust to several endogeneity controls.

From a theoretical point of view, there are reasons to expect that product innovation and not process innovation drives firms into exporting. Klepper (1996) analyzes the patterns of exit, entry, innovation and growth over the product life cycle. His findings indicate that firms are more likely to conduct product innovations in the beginning of their life cycle (prior to exporting), while they are more likely to focus on process innovations during the later stages of their life cycle. This pattern is in line with the product life cycle as put forth by Vernon (1966), where firms first introduce a product innovation on the domestic market, after which they start exporting that product. Rationalization of the production process, for instance through process innovations aimed at improving the efficiency of production only takes place at a later stage. As noted by Cassiman and Martinez-Ros (2007), process innovation is also likely to become more attractive to the firm once production volumes are large and competition is mounting.

However, to the extent that the introduction of a process innovation makes the firm more productive, process innovations can help firms to attain the minimum efficiency level needed to enter the export market in a profitable way.

3. Data and empirical facts

To investigate the relationship between a firm's innovation activities and its probability to start exporting, we use data from the Community Innovation Survey (CIS) for Belgium, obtained from the Belgian Science Policy (Belspo, 2006). The population for the CIS survey is selected on the basis of the full population of Belgian firms, registered at the National Office for Social Security at the end of the period considered (2000 for CIS3, 2004 for CIS4). Of these, all firms with at least ten employees are selected. Sampling is performed on the population after stratifying according to sector (NACE two-digit, three-digit in some cases), size (three size classes) and region (two-digit NUTS) (Teirlinck, 2005). The full sample of CIS3 firms contains 2,100 firms; while CIS4 has data on 3,322 firms. The data for CIS3 pertain to the years 1998-2000, while the CIS4-data is for the period 2002-2004.

The CIS questionnaire contains detailed information on firms' innovation activities, as well as some general information, such as firm's export intensity in 2000 (CIS3) and 2004 (CIS4). The survey has information on both innovative efforts of the firm (internal and external R&D) as well as on its innovative output. For innovative output, a distinction is made between a product innovation, defined as a new or significantly improved good or service that is new to the market or new to the firm; and process innovation, which concerns new or significantly improved methods of production, logistics, etc[7].

Apart from export intensity, the CIS questionnaire contains information on firm-level sales and employment and identifies foreign affiliates of multinational firms. In order to obtain additional information required to calculate the productivity of firms in the sample, we merge the CIS data with firm-level annual accounts information, obtained through the Belfirst database (BvDEP, 2006).

[7] For the definitions of all variables used in the empirical analysis, we refer to Appendix A.

As was already noted in the introduction, we restrict the sample used in the empirical analysis in two important ways. First, we limit attention to those firms that have replied to the questionnaire in two consecutive periods. Since sampling for the CIS survey is performed randomly, the overlap between the two periods is limited to 600 firms (i.e. these firms have responded to both questionnaires). Reducing the sample in this way allows us to use (four-year) lagged innovation and firm-level characteristics in the empirical analysis, and hence to avoid a simultaneity bias resulting from the fact that firm-level innovation and export decisions are taken at the same point in time.

Second, to control for past exporting history, we restrict the sample to two types of firms: i) firms that start exporting in 2004 (i.e. they did not export in 2000) and ii) a control group of firms that did not export in either period (Never exporters). 97 firms in the sample start exporting in 2004, while 92 firms did not export in 2000 or 2004. Hence, the total sample size amounts to 189 firms.

Table 1 summarizes the sector distribution specifically for our sample. As can be seen in the table, the sample covers all sectors of the economy. Apart from the number of firms, table 1 also lists the number of non-exporters and starters in each of the sectors considered.

[Table 1]

Ever since the seminal work by Bernard and Jensen (1995), many empirical papers have documented the differences between exporters and non-exporters in terms of several firm characteristics, such as size, productivity, etc. (see for instance De Loecker, 2007 for Slovenia or Mûuls and Pisu, 2009 for Belgium). In a similar vein, Table 2 reports summary statistics (mean and standard deviation) of a number of firm-level characteristics, separately for non-exporters and starters on the export market. However, unlike the papers cited above, Table 2 looks at the difference between exporters and non-exporters *prior* to their potential entry on the export market.

As was noted in the previous section, we will include lagged firm-level (innovation) characteristics in the empirical analysis in order to avoid a simultaneity bias, resulting from the fact that firms' innovation and export decisions, as well as decisions related to the allocation of inputs and outputs are taken at the same point in time. By including (four-year) lagged firm characteristics, we aim to control for these simultaneity issues. Analogously to the empirical analysis, Table 2 therefore reports lagged firm-level characteristics.

[Table 2]

Table 2 shows that exporters are larger and more productive[8] already four years prior to engaging on the export market. These differences are statistically significant. In the empirical analysis below, we take these differences into account, in addition to industry dummies to control for differences across sectors will be included.

Table 3 summarizes the innovation characteristics of the sample. Similar to Table 2, the table distinguishes between non-exporters and starters on the export market. The values reported in the table refer to the number of firms engaging in a particular innovation activity, the percentages are calculated with respect to the total number of non-exporters or starters, reported in the first row of the table. Several interesting facts emerge from Table 3.

[Table 3]

First, comparing the last two columns in Table 3, it is clear that firms that will start exporting in 2004, already exert greater innovative effort in 2000 compared to non-exporters in both periods. For internal and external R&D, the differences between the two groups are relatively small. About 30 percent of the starters engage in internal R&D in 2000, compared to 26 percent for the non-exporters. For external R&D, the relevant figures are 13 and 8 percent respectively. For innovative output however, the differences between the two groups are much larger. While 58 percent of the starters introduced a product innovation in 2000, only 33 percent of the non-exporters did. Similarly, 49 percent of the starters introduced a process innovation in 2000, compared to 26 percent for the non-exporters.

Second, as can be seen in the last row of Table 3, many firms introduce a product and process innovation simultaneously. Within the group of non-exporters, this is the case for 10 firms (accounting for about 11 percent of the number of non-exporters), while for the starters on the export market, this is true for 32 firms (or 33 percent of the number of starters). Hence, it is clear that firms, and particularly those firms that will start exporting in 2004, often carry out product and process innovations simultaneously rather than in isolation. Within the group of starters, 57 percent of all firms that introduced a product innovation simultaneously introduced a process innovation and 67 percent of the firms engaging in process innovation simultaneously engaged in product innovation. For the group of non-exporters, the relevant percentages are 33 and 42 percent respectively. The correlation between the two variables amounts to 0.4428. The overlap between these two different types of innovation will be taken into account in the empirical analysis below.

---

[8] Firm size is defined using employment data. Similar to Aw et al. (2007), total factor productivity is calculated using the index number methodology. While this methodology has a number of drawbacks, i.e. constant returns to scale and perfect competition are assumed and no allowance is made for unobservable factors, unlike parametric estimation, it does not assume a homogeneous production technology for all firms in a particular sector (Van Biesebroeck, 2007). Variables are defined in Appendix A.

4. Empirical model

In order to investigate to what extent firm-level innovation activities increase firms' export propensity, we estimate the following empirical model:

$$\Pr\left(START_{it}=1\right)=f\left[\ln\left(Size_{it-4}\right),\ln\left(TFP_{it-4}\right),INN_{it-4},I_i\right] \qquad [1]$$

where

$Size_{it-4}$	Firm-level employment in 2000;
TFP_{it-4}	Total factor productivity in 2000;
INN_{it-4}	Innovation characteristic, differs depending on specification;
I_i	Sector dummy.

The dependent variable in [1] is equal to one if the firm starts exporting in 2004 and zero otherwise. As noted before, our sample is limited to those firms that start exporting in 2004 and firms that did not export in both periods. Since we only have access to two consecutive CIS questionnaires, the use of initial characteristics in [1] implies that we can only include one year of data in the regression (2004). The year 2000 is used to define the lagged characteristics. For the definitions of the variables used in the empirical analysis, we refer to Appendix A.

We will include both innovative input and output measures in [1]. All innovation measures are defined as dummy variables, indicating whether the firm has engaged in a particular activity or not. We use two input indicators, referring to whether the firm has engaged in internal or external R&D in 2000[9] and two output indicators, referring to whether the firm has introduced a product or process innovation in 2000. As is illustrated in Table 4, the correlations between the different innovation variables are generally high. Only the correlations between the two output measures and external R&D are lower than 0.40, in all other cases, the values are larger than 0.40. As argued before, we will take this high correlation into account in the empirical analysis.

[Table 4]

Specifically, in order to avoid multicollinearity issues, which might result in the insignificance of some of the variables caused by the high correlation between them, we

[9] Aw et al. (2007) note that R&D intensity, unlike the discrete choice to engage in R&D, is more likely to be driven by firm-specific unobservable factors and noise. Therefore, although the CIS data report data on firm-level expenditures on internal and external R&D, we follow Aw et al. (2007) and rely on dummy variables, indicating whether the firm has actively engaged in R&D, to measure firms' innovation activities. This choice is consistent with the existence of a fixed setup cost of R&D, such that once this cost has been incurred, the amount of R&D actually spent matters less.

include only one innovation measure at a time. Moreover, to take the large degree of overlap between product and process innovation into account, we will further distinguish between firms that have only introduced a product innovation, only a process innovation or both simultaneously. This will allow us to investigate to what extent the simultaneous introduction of a product and process innovation offers an advantage to the firm in terms of its export market prospects. In what follows, the results of the baseline specification given by [1] will be discussed.

5. Empirical results

Table 5 reports the regression results for the baseline specification given by equation [1]. All regressions in the table include a full set of industry fixed effects[10] to control for differences across sectors. Each of the four columns in the table includes a different innovation measure. In the first two columns, input measures are added, while the last two columns report results using innovation output measures. All the values reported in Table 5 are marginal effects, defined as the marginal probability change at the mean of the independent variables (discrete change from 0 to 1 for dummy variables), standard errors are reported between brackets.

[Table 5]

Results in Table 5 show that productivity has a positive and significant influence on firms' propensity to start exporting. This result is in line with the theoretical and empirical self-selection literature (Melitz, 2003; Muûls and Pisu, 2009), i.e. only the more productive firms are able to enter the export market. Although Table 2 indicated that firms that start exporting are (on average) larger than their non-exporting counterparts, firm size is only (marginally) significant in Table 5. These results suggest that the differences between starters and non-exporters in terms of their size is mainly due to differences across sectors and not so much to differences within a sector.

Furthermore, in line with the results obtained by Aw et al. (2007) and Cassiman and Martinez-Ros (2007), our results suggest that firm-level investments in R&D (internal or external) do not result in a higher propensity to export in the next period. The last two columns of Table 5 show the results of estimating [1], but now including innovation output rather than input measures. We include product and process innovation separately here, without taking into account that many firms introduce both innovations simultaneously. The results suggest that both product and process innovation (irrespective of whether they were introduced in isolation or simultaneously) have a significantly positive impact on firms' propensity to start exporting.

[10] Sectors are grouped as in Table 1.

Specifically, the magnitude of the marginal effects implies that firms that introduce a product innovation increase their probability to start exporting by 22 percentage points, compared to 19 percentage points for process innovation. These findings are in line with those reported by Caldera (2009) for Spain. Using a similar empirical framework[11], she finds that firms introducing a product innovation increase their export propensity by 16 percentage points. Firms introducing a process innovation exhibit a 7 percentage points increase in their probability to export, which is somewhat lower than in our case.

To determine to what extent the correlation between product and process innovation leads to serious multicollinearity issues, the first column of Table 6 reports the results of the baseline specification, which now includes both innovation output variables, i.e. product and process innovation are both added as independent variables in the regression. When both innovation variables are taken into account simultaneously, only product innovation emerges as a significant determinant of firms' export propensity. These results are in line with results reported by Cassiman and Martinez-Ros (2007) who also report a positive and significant effect for product innovation, but not for process innovation on firms' export propensity.

[Table 6]

However, given the high correlation between the two innovation output measures and the fact that they both act as significant drivers of firms' probability to enter the export market, it can be argued that the insignificance of the process innovation variable does not reflect its true impact. Moreover, while including the innovation measures one by one avoids the multicollinearity issues discussed above, it fails to take into account potential complementarities between firms' product and process innovation in shaping their future export prospects. As was already noted in Section 3, 49 percent of all firms that introduced a product innovation in 2000 simultaneously introduced a process innovation. Similarly, 58 percent of all process innovators were also product innovators in 2000.

To take this high correlation into account, Table 6 distinguishes between four types of firms: i) non-innovators (the baseline), ii) firms that only introduced a product innovation in 2000, iii) firms that only introduced a process innovation in 2000 and iv) firms that introduced both a product and process innovation simultaneously. Since these categories are mutually exclusive (a firm is never part of more than one of the four groups), we avoid potential multicollinearity issues. Moreover, by accounting explicitly for the fact that some firms introduce a product and process innovation at the same time, we are able to determine to what extent both innovation activities have complementary effects on firms' export propensity.

[11] Caldera additonally adds random effects to the baseline specification. Since we only have data for two time periods and we add lagged firm characteristics in the regression, we cannot estimate a random effects probit model.

Results of the baseline model, but now including three rather than two innovation output measures, are reported in the second column of Table 6. Again, all regressions include sector dummies. Similar to the results for the non-innovation characteristics reported in Table 5, total factor productivity emerges as a significant driver of firms' export propensity, while firm size is insignificant. . For the innovation measures, results suggest that it is the simultaneous introduction of a product and process innovation, and not so much either of the two in isolation, that drives firms into exporting. Firms introducing a product or process innovation in isolation, exhibit no significant increase in their probability to start exporting.

This finding is in line with findings of Becker and Egger (2007) for Germany, who also find that the simultaneous introduction of a process and product innovation has a large impact on firms' export propensity. However, while Becker and Egger (2007) additionally find a positive and significant impact of product innovation in isolation (though not for process innovation), this is not the case here. Product or process innovations conducted in isolation exert no significant impact on the probability of firms to start exporting.

6. Accounting for anticipation effects

If firms can anticipate entry on the export market and their innovation activities are driven by this prospect[12], innovation cannot be considered exogenous in the analysis reported above. To control for this potential endogeneity, we will report several instrumental variable estimations for the innovation output measures. We choose to rely on two-stage least squares (instrumental variables or IV) regression to estimate the causal impact of firm-level innovation activities on its export propensity for two reasons. First, unlike linear IV models, non-linear IV estimation requires fairly strong assumptions, i.e. the error terms in the first and second stage need to be identically normally distributed and both stages need to be correctly specified for consistent estimation (Carrasco, 1998). Moreover, standard IV probit estimation procedures[13] require the endogenous variable to be continuous (i.e. the first estimation stage is linear), yielding inconsistent standard errors for endogenous dummy variables. We therefore follow Caldera (2009) and rely on two-stage least squares regression to investigate the causal impact of firm-level innovation activities on its export propensity.

As a first step, we estimate the preferred model of Table 6 (column II), including the three dummies representing whether the firm introduced a product or process innovation in isolation or the two of them simultaneously, but now using a Linear Probability Model (LPM). This will allow us to determine to what extent the LPM results are comparable to the probit results reported in Table 6. The results for the innovation measures are similar to the ones obtained with the probit model. Again we find that only those firms that introduce a

[12] The prospect of export market entry can be driven by anticipated trade liberalization as in Costantini and Melitz (2007) or by firm-level considerations.

[13] For instance, in Stata 10, IV Probit estimation can be achieved using the *ivprobit* command if the endogenous regressors are continuous.

product and process innovation simultaneously exhibit a significant increase in their probability to enter the export market.

While the coefficient on productivity is lower for the Linear Probability Model in Table 7, it is still positive and significant. The next three columns of Table 7 report results of applying an instrumental variables approach (IV) in the LPM. We account for the endogeneity of firms' innovation activities by instrumenting. Generally, instruments need to satisfy two requirements (Greene, 2008). First, they cannot have a direct impact on the dependent variable (i.e. on the probability to start exporting). Second, they need to be correlated with the endogenous regressor, conditional on all other covariates. Since there are three endogenous regressors in Table 6, we need at least three instruments.

The insignificance of the internal and external R&D dummy in Table 5 (i.e. they have no direct impact on the probability to start exporting), combined with the fact that internal and external R&D are essentially the inputs for the innovation outcomes (the endogenous variables), suggests they might be good instruments. Additionally, it is likely that firm-level on-the-job training activities, on which we have information from the Belfirst database (BvDEP, 2006) are correlated with firm-level innovation activities and in particular process innovation, since new production processes need to be executed and therefore introduced to employees and workers. While firm-level training (which is measured using a dummy variable) does not feature in Table 5 and 6, we ran an auxiliary regression[14] to ensure that training is not directly related to firms' propensity to start exporting.

To investigate to what extent the instruments are sufficiently "strong", i.e. are correlated with the endogenous dummy regressors conditional on all other covariates, we estimate the baseline model of Table 6 (column II) using the three instruments above. The first-stage results of the estimation procedure are reported in the Appendix (Table A.1)[15]. From Table A.1, it can be seen that for each of the three endogenous dummies (Only production innovation, Only process innovation and Both) at least one of the instruments yields a positive and significant coefficient. These results confirm our prior that the instruments chosen are indeed correlated with our endogenous regressors, conditional upon all other covariates.

The last four columns in Table 7 show the results for the innovation output measures, after accounting for potential endogeneity of firms' innovation activities (i.e. the anticipation effect). Similarly to Table 6, we distinguish between firms that have introduced a product or process innovation in isolation and those that have introduced both of them together. Surprisingly, both size and productivity are insignificant in all three columns.

Results in the last three columns of Table 7 suggest that, after accounting for the potential endogeneity of the innovation decision, firm-level innovation has no significant impact on firms' export propensity. While these results are not in line with those of Caldera (2009) and Cassiman and Martinez-Ros (2007) that both report a positive and significant impact of firms'

[14] Unreported, but available from the authors upon request.

[15] The second-stage results are reported in Table 7, last column.

innovation activities on its export propensity after accounting for the potential endogeneity of the innovation measures; they are in line with results reported by Damijan et al. (2008) for Slovenia, who fail to find a significant effect of firm-level innovation on the probability of firms to enter the export market.

Hence, after controlling for potential endogeneity of the innovation activities in [1], we find no evidence that firms engaging in product and/or process innovation are more likely to start exporting. These results suggest that only firms with a sufficiently high probability to start exporting will engage in product and process innovation prior to their entry on the export market, pointing to the importance of *self-selection into innovation activities*.

To test the validity and strength of our instruments, three test statistics are reported in Table 7. All test statistics are obtained using the Stata module *ivreg2*, developed by Baum, Schaffer and Stillman (2004). The Sargan-Hansen statistic tests for over-identification of the model, failure to reject the null hypothesis that the model is over-identified indicates that the instruments are valid. The Kleibergen-Paap statistic on the other hand tests for under-identification of the model by testing whether the model is of full rank. The null hypothesis states that the model is under-identified, rejection of the null implies that the model is identified. Finally, the Anderson-Rubin F-statistic tests whether the first-stage regressors are jointly significant and whether the model is identified. The Anderson-Rubin test is robust to the presence of weak instruments. Failure to reject the null hypothesis that the model is identified indicates that the instruments are valid.

Apart from the last column of Table 7, all test statistics indicate that the instruments used are indeed valid. The Sargan-Hansen test is never significant[16], suggesting that the model is correctly specified. The Kleibergen-Paap test statistic rejects the null hypothesis of under-identification at the five percent level in all but the last column of Table 7. Finally, the Anderson-Rubin F-statistic is never significant, suggesting that the model is identified and the instruments are valid. However, it is worth noting that the Kleibergen-Paap test statistic points to potential under-identification of the model in the last column of Table 7, where all three endogenous regressors are included together in the model. Although the other identification tests do not confirm this result, some caution in the interpretation of our result is warranted. Future research, ideally based on both a larger sample and including a time dimension, needs to be undertaken to confirm these results.

7. Conclusion

[16] The Sargan-Hansen test requires the model to be over-identified, i.e. there should be more instruments than endogenous variables. This implies that the test statistic can not be calculated for the last column of Table 7, where we have three endogenous regressors and three instruments.

This paper has explored the relationship between firm-level innovation activities and firms' propensity to start exporting, using data from the Community Innovation Survey for Belgium in two consecutive periods. The analysis fits in with a small, but growing body of literature where firm productivity is considered to be endogenous, rather than the result of an exogenous draw as in earlier models (e.g. Melitz, 2003). One of the ways in which firms can increase their productivity is through firm-level innovation or technology adoption (Bustos, 2005; Yeaple, 2005).

Several recent papers have explored the link between the innovation activities of firms and their propensity to start exporting empirically (e.g. Becker and Egger, 2007; Caldera, 2009; Cassiman and Martinez-Ros, 2007; Damijan et al., 2008). However, thus far, empirical results have been mixed and seem to depend on the type of innovation measures used, countries analyzed and methodologies used.

In our empirical analysis, we control for three potential sources of endogeneity: (i) simultaneity, which is a consequence of the simultaneous character of innovation and export decisions; (ii) causality, introduced by persistence in exporting activities and (iii) anticipation, caused by the fact that firms might innovate ahead of export market entry if their future export prospects are good. We account for these sources of bias by using lagged firm-level and innovation characteristics, by focusing on starters on the export market (versus a control group of non-exporters) and by applying instrumental variable estimation. A central finding of the analysis is that it is important to take the potential complementarities between product and process innovation into account when analyzing firms' propensity to export. Taking into account that about half of all innovating firms introduce a product and process innovation simultaneously, our empirical results suggest (before taking into account potential anticipation effects) that it is the combination of product and process innovation, rather than either of the two in isolation, that drives firms into the export market.

Furthermore, results point to the importance of taking the anticipation effect into account. After applying instrumental variables estimation, our results point to the importance of self-selection into product and/or process innovation: only those firms that have good prospects of entering the export market in the next period are more likely to invest in innovation activities. However, given the limitations inherent in the data used in the empirical analysis (no time dimension and a relatively small sample), future work in this area is needed to confirm these results.

References

Alvarez, Roberto and Lopez, Ricardo A. "Exporting and Performance: Evidence from Chilean Plants." *Canadian Journal of Economics*, 2005, *38*(4), pp. 1384-400.

Aw, Bee Yan; Roberts, Mark J. and Winston, Tor. "Export Market Participation, Investments in R&D and Worker Training, and the Evolution of Firm Productivity." *The World Economy*, 2007, *30*(1), pp. 83-104.

Baum, Christopher F.; Schaffer, M.E. and Stillman, Steven. "Instrumental Variables and GMM: Estimation and Testing." *Stata Journal*, 2004, *4*(2), pp. 1-31.

Becker, Sascha O. and Egger, Peter. "Endogenous Product Versus Process Innovation and a Firm's Propensity to Export." *CESIFO Working Paper Series*, 2007, *1906*.

Belspo. *Cis Questionnaire*. Brussels: Belgian Science Policy, 2006.

Bernard, Andrew B. and Jensen, Bradford J. "Exporters, Jobs and Wages in Us Manufacturing: 1976-1987." *Brookings Papers on Economic Activity - Microeconomics*, 1995, *1995*, pp. 67-119.

Bernard, Andrew B. and Jensen, J. Bradford. "Exceptional Exporter Performance: Cause, Effect or Both?" *Journal of International Economics*, 1999, *47*(1), pp. 1-25.

Bustos, Paula. "The Impact of Trade on Technology and Skill Upgrading: Evidence from Argentina." *Harvard University, mimeo.*, 2005.

BvDEP. *Belfirst Database*. DVD-rom. Bureau van Dijk Electronic Publishing, 2006.

Caldera, Aida. "Innovation and Exporting: Evidence from Spanish Manufacturing Firms." *ECARES Working Paper Series*, 2009, *2009-014*.

Carrasco, R. "Binary Choice with Binary Endogenous Regressors in Panel Data: Estimating the Effect of Fertility on Female Labour Participation" Journal of Business and Economics Statistics 2001, 19(4), pp. 385-94.

Cassiman, Bruno and Martinez-Ros, Ester. "Product Innovation and Exports: Evidence from Spanish Manufacturing." *unpublished manuscript*, 2007.

Castellani, Davide and Zanfei, Antonello. *Multinational Firms, Innovation and Productivity*. Cheltenham: Edward Elgar Publishing, 2006.

Castellani, Davide and Zanfei, Antonello. "Internationalisation, Innovation and Productivity: How Do Firms Differ in Italy?" *The World Economy*, 2007, *30*(1), pp. 156-76.

Caves, Douglas W.; Christensen, Laurits R. and Diewert, W. Erwin. "Multilateral Comparisons of Output, Input, and Productivity Using Superlative Index Numbers." *The economic journal*, 1982, *92*(365), pp. 73-86.

Costantini, James A. and Melitz, Marc. "The Dynamics of Firm-Level Adjustment to Trade Liberalization." *Princeton University, mimeo*, 2007.

Criscuolo, Chiara; Haskel, Jonathan E. and Slaughter, Matthew J. "Global Engagement and the Innovation Activities of Firms." *NBER Working Paper Series*, 2005, (11479), pp. 48.

Damijan, Joze P.; Kostevc, Crt and Polanec, Saso. "From Innovation to Exporting or Vice Versa? Causal Link between Innovation Activity and Exporting in Slovenian Microdata." *Licos Discussion Paper Series*, 2008, *204*.

De Loecker, Jan. "Do Exports Generate Higher Productivity? Evidence from Slovenia." *Journal of International Economics*, 2007, *73*(1), pp. 69-98.

Greenaway, David and Kneller, Richard. "Firm Heterogeneity, Exporting and Foreign Direct Investment." *The Economic Journal*, 2007, *117*(517), pp. F134-F61.

Greene, William H. *Econometric Analysis*. New Jersey: Prentice Hall, 2008.

Griffith, Rachel; Huergo, Elena; Mairesse, Jacques and Peters, Bettina. "Innovation and Productivity across Four European Countries." *Oxford Review of Economic Policy*, 2006, *22*(4), pp. 483-98.

Hopenhayn, Hugo A. "Entry, Exit and Firm Dynamics in Long Run Equilibrium." *Econometrica*, 1992, *60*(5), pp. 1127-50.

Iacovone, Leonardo and Smarzynska Javorcik, Beata. "Shipping Good Tequila Out: Investment, Domestic Unit Values and Entry of Multi-Product Plants into Export Markets." *Unpublished manuscript*, 2008.

Jovanovic, Boyan. "Selection and the Evolution of Industry." *Econometrica*, 1982, *50*(3), pp. 649-70.

Klepper, Steven. "Entry, Exit, Growth, and Innovation over the Product Life Cycle." *The American Economic Review*, 1996, *86*(3), pp. 562-83.

Mairesse, Jacques and Mohnen, Pierre. "Accounting for Innovation and Measuring Innovativeness: An Illustrative Framework and an Application." *American Economic Review*, 2002, *92*(2), pp. 226-30.

____. "The Importance of R&D for Innovation: A Reassessment Using French Survey Data." *The Journal of Technology Transfer*, 2004, *30*(1), pp. 183-97.

Melitz, Marc J. "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity." *Econometrica*, 2003, *71*(6), pp. 1695-725.

Muûls, Mirabelle and Pisu, Mauro. "Imports and Exports at the Level of the Firm: Evidence from Belgium." *World Economy*, 2009, *32*(5), pp. 692-734.

Teirlinck, Peter. "Report CIS 4," *Belgian Science Policy.* 2005.

Van Biesebroeck, Johannes. "Robustness of Productivity Estimates " *Journal of Industrial Economics*, 2007, *55*(3), pp. 529-69.

Vernon, Raymond. "International Investment and International Trade in the Product Cycle." *The Quarterly Journal of Economics*, 1966, *80*(2), pp. 190-207.

Wagner, Joachim. "Exports and Productivity: A Survey of the Evidence from Firm-Level Data." *The World Economy*, 2007, *30*(1), pp. 60-82.

Yeaple, Stephen Ross. "A Simple Model of Firm Heterogeneity, International Trade, and Wages." *Journal of International Economics*, 2005, *65*(1), pp. 1-20.

Appendix A: Data and definition of variables.

All innovation variables and information on firms' export status were obtained from the CIS questionnaire (Belspo, 2006). All accounting data are obtained from the Belfirst database (BvDEP, 2006). The definitions of the variables are given below, capital letters refer to dummy variables.

Dependent variable

$START_{it}$	dummy equal to one if the firm starts exporting in 2004 (no exports in 2000).

Independent variables

$Size_{it-4}$	Firm-level employment in 2000, expressed in full-time equivalents.
TFP_{it-4}	Total factor productivity in 2000, defined using index numbers (see below).
INN_{it-4}	Innovation dummy. Seven different innovation dummies are used in the empirical analysis, see below for their definitions.

Total factor productivity

To obtain comparable levels of total factor productivity (TFP) across firms, we follow Aw et al. (2007) and apply the Caves, Christensen and Diewert (1982) methodology. Specifically, applying this methodology to a value added production function yields the following formula to calculate comparable levels of TFP across firms.

$$\ln \tilde{A}_i^{VA} = \ln A_i^{VA} - \overline{\ln A^{VA}} = \left(\ln Q_i - \overline{\ln Q}\right) - \tilde{s}_i^L \left(\ln L_i - \overline{\ln L}\right) - \tilde{s}_i^K \left(\ln K_i - \overline{\ln K}\right) \qquad [A.1]$$

where $\tilde{A}_i^{VA}$ refers to the total factor productivity index, bars over variables indicate sample means; Q_i, L_i and K_i stand for output, labour and capital respectively and s refers to factor shares, which are defined as follows.

$$\tilde{s}_i^L + \tilde{s}_i^K = 1$$

$$\tilde{s}_i^X = \frac{s_i^X + \overline{s^X}}{2} \text{ where } \overline{s^X} = \frac{1}{N}\sum_{i=1}^{N} s_i^X \text{ , } s_i^X = \frac{X_i P_i^X}{Q_i P_i} \text{ and } X = L, K$$

The index calculated on the basis of [A.1] is a Törnqvist-Theil-translog index. Intuitively, the index is calculated by comparing each firm to a hypothetical firm, where the hypothetical firm is defined as the average over all firms as illustrated above.

A number of assumptions are imposed when TFP is calculated according to [A.1] (Van Biesebroeck, 2007): 1) perfect competition in output and input markets, 2) firms are profit-maximizing agents, 3) no measurement error and 4) constant returns to scale. The last of these assumptions can be relaxed if outside information on the extent of economies of scale is available to the researcher. Important advantages of the index number methodology are that it can readily be implemented and that it allows for heterogeneity in production technology across firms. Disadvantages associated with index numbers are its deterministic nature and the imposed assumptions on market structure and firm behavior (Van Biesebroeck, 2007).

Innovation measures

INTERNAL R&D DUMMY

Dummy equal to one if the firm engaged in internal R&D activities in 2000. Internal R&D activities are defined as "creative work undertaken within your enterprise to increase the stock of knowledge and its use to devise new and improved products and processes (including software development)" in the CIS questionnaire.

EXTERNAL R&D DUMMY

Dummy equal to one if the firm engaged in external R&D activities in 2000. External R&D activities are defined as "Extramural R&D: same activities as above, but performed by other companies (including other enterprises within your group) or by public or private research organisations and purchased by your enterprise." in the CIS questionnaire.

PRODUCT INNOVATION DUMMY

Dummy equal to one if the firm introduced a product innovation in 2000. A product innovation is defined as follows in the CIS questionnaire: "New or significantly improved goods or services (Exclude the simple resale of new goods purchased from other enterprises and changes of a solely aesthetic nature.)."

PROCESS INNOVATION DUMMY

Dummy equal to one if the firm introduced a process innovation in 2000. A process innovation is defined as follows in the CIS questionnaire: "i) New or significantly improved methods of manufacturing or producing goods or services, ii) New or significantly improved logistics, delivery or distribution methods for your inputs, goods or services or New or significantly improved supporting activities for your processes, such as maintenance systems or operations for purchasing, accounting or computing."

ONLY PRODUCT INNOVATION

Dummy equal to one if the firm introduced a product innovation in 2000, but no process innovation.

ONLY PROCESS INNOVATION

Dummy equal to one if the firm introduced a process innovation in 2000, but no product innovation.

PRODUCT AND PROCESS INNOVATION
Dummy equal to one if the firm introduced a process *and* product innovation in 2000.

Instruments

Innovation measure	Instrument
ONLY PRODUCT INN	Internal R&D dummy (firm-level) in 2000 External R&D dummy (firm-level) in 2000 Training dummy (firm-level) in 2000
ONLY PROCESS INN	Internal R&D dummy (firm-level) in 2000 External R&D dummy (firm-level) in 2000 Training dummy (firm-level) in 2000
ONLY PRODUCT INN ONLY PROCESS INN PRODUCT AND PROCESS	Internal R&D dummy (firm-level) in 2000 External R&D dummy (firm-level) in 2000 Training dummy (firm-level) in 2000

Table A.1. First-stage regression results			
Dependent variable	Onlyprod2000	Onlyproc2000	Bothinn2000
Size	-0.005	-0.04	0.067**
ln(employment) in 2000	[0.035]	[0.028]	[0.030]
Total factor productivity	-0.001	0.001	-0.001
Törnqvist index, logs 2000	[0.001]	[0.001]	[0.001]
Internal R&D	**0.231****	**0.213****	**0.143***
(Firm-level dummy, 2000)	**[0.100]**	**[0.103]**	**[0.083]**
External R&D	-0.105	-0.188	**0.318****
(Firm-level dummy, 2000)	[0.145]	[0.138]	**[0.125]**
Training	0.107	**0.153****	0.025
(Firm-level dummy, 2000)	[0.084]	**[0.068]**	[0.078]
Sector dummies	Yes	Yes	Yes
Number of observations	189	189	189

Results of first-stage regression of the IV estimation reported in the last column of Table 6. Reported values are coefficients [standard errors]. The dependent variables are the three innovation dummies (listed at the top of the column)Instruments are internal and external R&D dummy for 2000 and the training dummy, all observed a the *firm-level.* Significance level: *** p<0.01, ** p<0.05, * p<0.1.

Table 1: Sector distribution			
Sector	*N*	*Non-exporters*	*Starters*
Mining (Nace 14)	2	1	1
Food, beverages & tobacco (Nace 15-16)	6	1	5
Textiles, clothing, leather (Nace 17-19)	6	2	4
Wood (products) (Nace 20)	3	1	2
Paper and publishing (Nace 21-22)	5	1	4
Fuel and chemicals (Nace 23-24)	5	1	4
Rubber and plastics (Nace 25)	4	1	3
Non-metallic minerals (Nace 26)	4	1	3
Basic and fabricated metals, machinery (Nace 27-29)	24	4	20
Electrical, optical, medical instruments (Nace 30-33)	3	1	2
Tranport equipment, manufacturing n.e.c. (Nace 34-37)	9	6	3
Construction (Nace 45)	2	2	0
Wholesale and retail trade (Nace 50-52)	35	16	19
Transport and financial services (Nace 60-67)	40	30	10
Real estate and business services (Nace 70-74)	41	24	17
Total	**189**	**92**	**97**

Table 2: Comparing starters to non-exporters: Initial firm-level characteristics		
Variable	*Non-exporters*	*Starters*
Number of firms	92	97
[Percentage of total]	[48.68%]	[51.32%]
Size	3.65	3.98***
(Employment, fte, 2000)	[1.28]	[1.32]
Total factor productivity	1.00	2.25*
(Törnqvist productivity index, 2000)	[0.49]	[9.11]

Reported values are means [standard deviations] in 2000 (except where the number of firms is reported). Starters are firms that start exporting in 2004, non-exporters do not export in 2000 and 2004. Significance levels (*** $p < 0.01$; ** $p < 0.05$; * $p < 0.10$) refer to one tailed test on the difference between the means for the starters compared to non-exporters. Variables are defined in Appendix A.

Table 3: Comparing starters to non-exporters: Innovation characteristics		
Variable	*Non-exporters*	*Starters*
N	92 [48.68%]	97 [51.32%]
Number of firms engaging in internal R&D	24 [26.09%]	29 [29.90%]
Number of firms engaging in external R&D	7 [7.61%]	13 [13.40%]
No. of firms introducing a product inn.	30 [32.61%]	56 [57.73%]
No. of firms introducing a process inn.	24 [26.09%]	48 [49.48%]
No. of firms introd. product and process inn.	10 [10.87%]	32 [32.99%]

Reported values are number of observations [percentage of total number of firms in that column]. Starters are firms that start exporting in 2004, non-exporters do not export in 2000 and 2004. Variables are defined in Appendix A.

Table 4: Correlations innovation measures				
	[1]	[2]	[3]	[4]
Internal R&D dummy	1			
External R&D dummy	0.4745	1		
Product innovation dummy	0.5176	0.3074	1	
Process innovation dummy	0.4805	0.3323	0.4428	1

Table 5: Regression results				
	Input measures		*Output measures*	
Variables	*Internal R&D dummy*	*External R&D dummy*	*Product innovation*	*Process innovation*
Size	0.051*	0.044	0.030	0.031
ln(Employment) in 2000	[0.031]	[0.031]	[0.031]	[0.031]
Total factor productivity	0.083**	0.077**	0.064*	0.077**
Törnqvist index, logs, 2000	[0.036]	[0.035]	[0.035]	[0.036]
Innovation measure	**-0.066**	**0.064**	**0.217*** **	**0.185** **
	[0.100]	**[0.131]**	**[0.082]**	**[0.087]**
Sector dummies	Yes	Yes	Yes	Yes
Number of observations	189	189	189	189
Pseudo R-square	0.166	0.165	0.190	0.182

Each column reports the results of a probit regression, where the dependent variable is the probability to start exporting in 2004. Each column includes a different innovation dummy as independent variable (listed at the top of the column), in addition to size, productivity and sector dummies. Reported values are marginal effects [standard errors], defined as the marginal probability change at the mean of the independent variable or the discrete change of a dummy variable from 0 to 1. Significance level: *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Variables are defined in Appendix A.

Table 6: Regression results					
Variables	*I*	*II*	*III*	*IV*	*V*
Size	0.023	0.018	0.045	0.047	0.022
ln(Employment) in 2000	[0.032]	[0.033]	[0.030]	[0.030]	[0.032]
Total factor productivity	0.066*	0.068*	0.075**	0.077**	0.073**
Törnqvist index, logs, 2000	[0.036]	[0.036]	[0.035]	[0.035]	[0.036]
Product innovation	**0.175****	**-**	**-**	**-**	**-**
(dummy, 2000)	**[0.089]**				
Process innovation	**0.119**	**-**	**-**	**-**	**-**
(dummy, 2000)	**[0.096]**				
Only product innovation		**0.101**	**0.085**	**-**	**-**
(dummy, 2000)		**[0.101]**	**[0.097]**		
Only process innovation		**0.029**	**-**	**-0.028**	**-**
(dummy, 2000)		**[0.123]**		**[0.116]**	
Product & Process innovation		**0.301*****	**-**	**-**	**0.285*****
(dummy, 2000)		**[0.086]**			**[0.085]**
Sector dummies	Yes	Yes	Yes	Yes	Yes
Number of observations	189	189	189	189	189
Psuedo R-square	0.196	0.201	0.167	0.165	0.196

Each column reports the results of a probit regression, where the dependent variable is the probability to start exporting in 2004. In addition to size, productivity and sector dummies, each regression includes a number of innovation dummies. Reported values are marginal effects [standard errors], defined as the marginal probability change at the mean of the independent variable or the discrete change of a dummy variable from 0 to 1. Significance level: *** p<0.01, ** p<0.05, * p<0.1. Variables are defined in Appendix A.

Table 7: Instrumental variables estimation					
Variables	Baseline model LPM	*Output measures*			
		IV	*IV*	*IV*	*IV*
Size	0.008	0.037	0.039	0.03	0.077
ln(Employment) in 2000	[0.028]	[0.028]	[0.025]	[0.036]	[0.270]
Total factor productivity	0.003**	0.002	0.002	0.002	-0.004
Törnqvist index, logs, 2000	[0.001]	[0.001]	[0.001]	[0.001]	[0.018]
Only product innovation	**0.107**	**0.065**	**-**	**-**	**-5.153**
	[0.096]	**[0.349]**			**[14.535]**
Only process innovation	**0.036**	**-**	**0.125**	**-**	**4.377**
	[0.120]		**[0.340]**		**[12.142]**
Product and process inn.	**0.275*****	**-**	**-**	**0.082**	**1.207**
	[0.092]			**[0.253]**	**[3.576]**
Instruments (dummies, 2000)	-	External R&D Internal R&D Training	External R&D Internal R&D Training	External R&D Internal R&D Training	External R&D Internal R&D Training
Sargan-Hansen overidentification test (Chi²)	-	3.437	3.328	3.279	-
Kleibergen-Paap underidentification test (Chi²)	-	7.775**	9.27**	11.26***	0.13
Anderson-Rubin weak instruments - robust inference test (F)	-	1.12	1.12	1.12	1.12
Sector dummies	Yes	Yes	Yes	Yes	Yes
R-square	0.236	0.198	0.184	0.212	0.178
Number of observations	189	189	189	189	189

With the exception of the first column, each column reports the results of an Instrumental Variables regression, where the dependent variable is a dummy equal to one if the firm starts exporting in 2004. Each column includes different innovation dummies as independent variables (listed at the top of the column), in addition to firm size, productivity and sector dummies. The innovation variables are instrumented using the variables listed. The first column reports OLS results for the baseline model (column II in Table 6), no variables are instrumented. Reported values are coefficients [standard errors]. Significance level: *** p<0.01, ** p<0.05, * p<0.1. Variables are defined in Appendix A. The null-hypothesis of the Sargan-Hansen is that the instruments are valid. For the Kleibergen-Paap test, the null-hypothesis is that the model is underidentified (not of full rank). The null hypothesis of the Anderson-Rubin test is that the variables in the first stage are jointly significant and that the model is identified. The Sargan-Hansen test can only be performed when the model is over-identified, i.e. there are more instruments than endogenous regressors.