Endogeneity between Internationalization and Knowledge creation of global R&D leader firms: An econometric approach using Scoreboard data

Areti Gkypali,[a] Kostas Tsekouras[a] and Nick von Tunzelmann[b]*

[a.] Department of Economics, University of Patras, Greece

[b.] Science & Technology Policy Research (SPRU), University of Sussex, UK

Abstract

Taking advantage of a dataset containing the world's leader firms in R&D expenditures, this paper is concerned with testing for the presence of endogeneity between the firm's knowledge creation process and the degree of internationalization, captured by R&D and export intensity respectively. Using econometric techniques of a simultaneous system of equations in a Tobit model and a simple regression on a pooled cross-section dataset, our results indicate that there exists an endogenous relationship between these two firm-level activities. In this endogenous framework we further analyze the interrelationship between R&D and export intensity. In addition we introduce into our analysis a rather neglected aspect of R&D intensity: that of the efficiency of the resources devoted to the knowledge creation process.

Keywords: export intensity, R&D intensity, knowledge creation, endogeneity, internationalization

*Corresponding Author:
SPRU (Science and Technology Policy Research)
Freeman Centre, University of Sussex
Falmer, Brighton, BN1 9QE
United Kingdom
Tel. +44 1273 873533
G.N.Von-Tunzelmann@sussex.ac.uk

Acknowledgements

The authors would like to thank Professor William Greene (Stern School of Business, NYU) for his valuable suggestions. We also thank Janet Butt from BIS (formerly DIUS) in the UK Government for her valuable help in providing the necessary information.

1. Introduction

Among the very large number of stimuli a firm has for expanding its knowledge base, conducting R&D and engaging in exporting activities are identified as two especially prominent activities that offer the firm opportunities in this direction (Johanson, 1977; Nonaka, 1994). In this paper we consider these two activities of firms as potential knowledge flows to augment their knowledge stock. Although their relationship has been investigated thoroughly by scholars (see references in section 2 below), the existence of degrees of endogeneity between these two activities tends to be missed by both empirical findings and theoretical predictions. Since internationalization is a vehicle for knowledge acquisition and diffusion on the global scale (Pedersen et al., 2003), it becomes highly relevant to examine if and how firms that in most cases have high status in global technological achievements are affected by knowledge creation in the context of their internationalization, with the aim of revealing any existing endogenous relationship between them.

In this paper we argue that R&D intensity captures some of the most significant aspects of the knowledge creation process while export intensity reflects one important dimension of the firm's degree of internationalization. Taking advantage of information provided by a UK government department regarding R&D leaders (RDL firms) we devised a pooled cross-section dataset for the years 2007-2008 and test for the presence of endogeneity with respect to R&D and export intensity. The econometric approach employed takes advantage of the intensity measures, instead of models of binary variables, exploiting the full information conveyed by the outcome of the RDL firms' decision-making processes with respect to their R&D and exporting activities.

Before doing so, attention should first be drawn to a rather neglected aspect of the R&D intensity, which may provide further valuable economic and managerial information. As will be argued in the current paper, R&D intensity is, through its definition as a ratio of inputs to outputs, an admittedly rough index of R&D process efficiency (Fare et al., 1994),[1] and thus of the efficiency of the knowledge creation process. With this rough approximation in mind, we therefore introduce an additional dimension into our analysis, namely that of the interrelationship between the degree of internationalization, as it is captured by export intensity, and the efficient allocation and exploitation of the resources devoted to knowledge creation.

The rest of the paper is organized as follows. In the next section a brief review of the literature regarding the empirical and theoretical findings of the relationship between R&D and exporting activities will be conducted. Section 3 presents some modelling issues and the econometric strategy to be followed, while section 4 provides the data and definition of variables. Section 5 is devoted to the discussion of the econometric results; section 6 concludes the paper and poses some further research directions.

2. Literature Review

2.1 A knowledge-based framework

The knowledge-based view of the firm (Grant, 1996) conceptualizes the firm as an entity that mainly creates knowledge in a unique way, which in turn constitutes its competitive advantage, and this is why the core competences of a firm are hard to imitate by competitors (Spender, 1996). Kogut and Zander (1992) argued that creating

[1] Such an R&D/Sales index follows an early notion of efficiency, as defined by Farrell (1957).

new knowledge does not occur in abstraction from current abilities, but rather new learning, such as innovation, is a product of the firm's combinative capabilities to generate new applications from existing knowledge. This view is helpful for exploring the processes by which a firm expands its knowledge base, along with the potential interdependence between these processes. In such a framework, R&D and exporting activities constitute opportunities serving this precise purpose, among others.

More specifically, knowledge may be acquired and recombined through internal and external learning. In-house R&D is one of the most critical means of creating new knowledge and is likely to result in a succession of product and process innovations (Cohen and Levinthal, 1989). In this direction, firms' exporting activities have been acknowledged by scholars (e.g. Saggi, 2002) as a channel for diffusing as well as transferring knowledge and technology. In addition, it has been argued that learning by exporting (Evenson and Westphal, 1995) is another way of acquiring knowledge. This in turn suggests that exporting activities offer opportunities for further exploitation and expansion of the knowledge base of the firm, through using them in conjunction with adaptation from the resources devoted to R&D.

2.2 Endogeneity and findings from the literature

The existence of an endogenous relationship between R&D and export intensity, even though recognised very early in the literature (e.g. Keesing, 1967), has only very recently begun to be investigated. More specifically, Clerides et al. (1998) apply causality tests in order to define the pattern of causality between R&D and export intensity and find that more productive firms choose to export. Smith et al. (2002), using a sample of Danish firms, tackle this issue of endogeneity and report that R&D increases the probability that a firm will become an exporting firm. Even more interesting is the empirical work by Harris and Li (2008): using a UK sample of firms, they investigate the endogeneity between R&D and export propensity and argue that a crucial factor for the lack of evidence on this endogenous relationship may be due to the lack of appropriate data and problems with econometric methods that allow testing for such an endogenous relationship.

Many scholars have been preoccupied with the economic logic underlying R&D and exporting activities. Do these measures signify increased productivity, are they a means for firm growth, are they a sign of increased competitiveness, etc.? Their relationship has been investigated and interpreted via theoretical frameworks such as the Product life cycle and Endogenous growth theory, as well as through empirical investigation.

More specifically, in the context of the Product life cycle theory it is argued that innovation will eventually lead to exporting. Scholars like Posner (1961), Vernon (1966), Krugman (1979) and Dollar (1986), among others, predicted that developed countries which are more R&D-intensive will be also more export-intensive due to the competitive advantage of the advanced technological capabilities they possess. On the other hand, the Endogenous growth models of international trade (Grossman and Helpman, 1989, 1990, 1991a; Segerstrom et al., 1990; Young, 1991; Aghion and Howitt, 1998, ch. 11) consider that innovative activity is endogenously determined and predict some dynamic effects from its relationship to international trade. Two possible explanations are provided as to why this may be the case. The first lies in the intense competition in international markets that imposes on exporting firms the responsibility for continuous improvement of their products and processes in order to remain competitive, thus increasing their probability of innovation, i.e. a demand-side stimulus. The second explanation is the so-called 'learning by exporting' phenomenon, in the sense that exporting firms' access to foreign markets provides them with feedback from their suppliers and/or customers, which gives them the opportunity to transform this knowledge into innovation, working through the supply side (Evenson and Westphal, 1995).

Empirical research on the relationship of export to R&D intensity has provided some mixed results. For instance, Schlegelmilch and Crook (1988) and Lefebvre et al. (1998) found that R&D intensity has no influence on export intensity. Lall and Kumar (1981), investigating a sample of Indian firms, revealed a negative relationship between export and R&D intensities. On the other hand, a positive and statistically significant relationship has been reported by several authors (Willmore, 1992; Wakelin, 1998; Wagner, 2001; Bleaney and Wakelin, 2002; Lachenmaier and Woessmann, 2006; Pla-Barber and Alegre, 2007).

A second task for this paper is the effort to incorporate the empirical findings regarding the determinants of R&D and export intensity in order to be able to gain a more holistic view of these firm-level activities. Despite the extensive literature investigating the determinants of R&D and exporting activities, there exists a difficulty in identifying any definite relationship between R&D intensity and the other possible determinants, owing to the difficulty of measurement itself, the lack of thorough and extensive data on R&D at the firm level, and the limited econometric techniques available at any point in time (Cohen and Levin, 1989). The same applies to the identification of the determinants of export intensity where, despite the substantial literature investigating the determinants of exporting activities, this field of research remains quite vague, in the sense that again the research on the identification of export intensity determinants has revealed controversial empirical evidence (van Dijk, 2002).

3. Modeling Issues

Let us consider the i^{th} RDL firm which is involved in two decision-making processes, namely the determination of its R&D activities $\left(RDINT\right)$ and its global orientation $\left(EXPINT\right)$ in period *t*. Considering the case where the knowledge intensity of the RDL firm is crucial for its internationalization level, the following structural equation describes the latter decision-making process:

$$EXPINT_{i,t} = RDINT_{i,t}\gamma_1 + \mathbf{z}'_{\mathbf{EXP}_{i,t}}\boldsymbol{\delta}_{\mathbf{EXP}} + \mathbf{u}_{\mathbf{EXP}_{i,t}} \qquad (1)$$

where $\mathbf{z}_{\mathbf{EXP}}$ is the vector of control variables capturing the variation of $EXPINT$ due to exogenous factors that affect the underlying decision-making process, $\boldsymbol{\delta}_{\mathbf{EXP}}$ the corresponding vector of parameters to be estimated and $\mathbf{u}_{\mathbf{EXP}}$ the error term. Parameter γ_1 depicts the influence of the RDL firm's knowledge creation intensity on its global orientation and is also going to be estimated. As regards the observability of the outcome of the decision-making process related to the RDL firm's level of internationalization the following rule applies:

$$EXPINT_{i,t} = \begin{cases} 0 \text{ if } EXPINT_{i,t} \leq 0 \\ EXPINT^{*}_{i,t} \text{ if } 0 < EXPINT_{i,t} \leq 1 \\ 1 \text{ if } EXPINT_{i,t} > 1 \end{cases} \qquad (2)$$

where $EXPINT^{*}_{i,t}$ is the observable value of export intensity of the i^{th} RDL firm in period t. That is, structural equation (1) is, by definition, a tobit equation.

On the other hand, and taking into account that the examined firms are R&D global leaders, a simple linear regression of the following form may describe the corresponding decision-making process regarding the level of their R&D intensity:

$$RDINT_{i,y} = \mathbf{z}'_{\mathbf{RD}_{i,t}}\boldsymbol{\delta}_{\mathbf{RD}} + \mathbf{u}_{\mathbf{RD}_{i,t}} \qquad (3)$$

where $\mathbf{z_{RD}}$ is the vector of the control variables capturing the factors engaged in the RDL firm's decision-making process with respect to the intensity of its R&D activities, $\boldsymbol{\delta}_{\mathbf{RD}}$ is the corresponding vector of parameters, and $\mathbf{u_{RD}}$ the error term with usual properties. Hereafter the subscripts i and t are suppressed for simplicity.

To summarize, (1) and (3) are the structural forms of equations that reflect the examined firm's decision-making processes regarding the determination of its level of internationalization, under the condition that the firm is a global R&D leader. In the case where the knowledge intensity of the firm is interrelated with its decision-making regarding its export intensity the following condition holds:

$$Cov\left(u_{EXP}, u_{RD}\right) \neq 0 \tag{4}$$

The validity of (4) involves the rejection of the exogeneity of the *RDINT* variable with respect to the *EXPINT* equation, or in other words the assumption that E(*RDINT*, $\mathbf{u_{EXP}}$) = 0 is violated. Thus, equations (1) and (3) form a system of simultaneous estimated equations where two characteristics are dominant: first, the recursive character of the system, and second the censored character of the *EXPINT* equation.

In order to cope with these special features we follow the approach introduced by Smith and Blundell (1986) closely. In particular, we assume that the correlation between the two error terms is of the following linear form:

$$u_{EXP} = u_{RD}^{'}\alpha + \varepsilon_{EXP} \tag{5}$$

with

$$\begin{bmatrix} u_{EXP} \\ u_{RD} \end{bmatrix} \sim NI\left(0, \begin{pmatrix} \sigma_{EXP}^{2} & \sigma_{RD,EXP}^{'} \\ \sigma_{RD,EXP} & \sigma_{RD}^{2} \end{pmatrix}\right) \tag{6}$$

Regarding the distribution of the ε_{EXP} term we assume that $\varepsilon_{EXP} \sim N\left(0, \sigma_{\varepsilon}^{2}\right)$ and it is independent of *RDINT* and u_{RD}. At this point it should be pointed out that σ_{RD}^{2} is computed taking into account that from the relation (5) the following holds:

$$\begin{aligned} Var\left(u_{RD}\right) &= Var\left(\frac{u_{EXP} - \varepsilon_{EXP}}{\alpha}\right) = \left(\frac{1}{\alpha^{2}}\right)\left[Var\left(u_{EXP} - \varepsilon_{EXP}\right)\right] \\ &= \left(\frac{1}{\alpha^{2}}\right)\left[Var\left(u_{EXP}\right) + Var\left(\varepsilon_{EXP}\right) - 2Cov\left(u_{EXP}, \varepsilon_{EXP}\right)\right] \end{aligned} \tag{7}$$

It is apparent that the variance of the error term of the *RDINT* equation depends on the parameter α and the variances of $u_{EXP}, \varepsilon_{EXP}$. The parameter α as reflected in relation (5) is the slope of the linear equation or in other words reflects the degree of correlation between the two error terms.

Even though we allow the structural form for R&D intensity $(RDINT)$ to depend directly on the latent variable for export intensity $(EXPINT^{*})$ it does not directly depend on the observable variable *EXPINT* (Heckman, 1978). Finally, in order to obtain a consistent estimator (with a known asymptotic normal distribution) for $\boldsymbol{\delta}_{\mathbf{RD}}$ and derive an estimator for α we finally estimate:

$$EXPINT^{*} = RD\hat{I}NT\gamma_{1} + \mathbf{z}_{\mathbf{EXP}}^{'}\boldsymbol{\delta}_{\mathbf{exp}} + \hat{u}_{\mathbf{RDINT}}\left(\alpha + \gamma_{1}\right) + \varepsilon_{\mathbf{EXP}} \tag{8}$$

where $RD\hat{I}NT$ is the estimated dependent variable of the OLS regression and $\hat{u}_{RD}$ are the residuals taken from the second estimated equation. Following Greene (2007) we calculated the parameter $\psi = \frac{\sigma_{12}}{\sigma_{2}^{2}}$ and tested for weak exogeneity using a simple t-test of the hypothesis that $\psi = 0$, i.e. $Corr\left(u_{EXP}, u_{RD}\right) = 0$, or $\alpha = 0$.[2]

[2] For further elaboration of the joint log-likelihood function along with a test for weak exogeneity, see the paper of Smith and Blundell (1986).

4. Sample and Variable Definitions

The constructed dataset is drawn from the R&D Scoreboard provided by the former Department of Innovation, Universities and Skills (DIUS) – now the Department for Business, Innovation and Skills (BIS) – and covers two time periods. More specifically, the data represents firms' financial and other basic economic characteristics for the years 2006 and 2007. Taking advantage of the R&D Scoreboard regarding years 2005 and 2004 our data set has been enriched with additional information.

Attention should be drawn to the fact that the reporting firms were selected on the basis of their R&D expenditures that are funded by themselves. R&D undertaken under contract for other agents such as governments or other companies, as well as the firm's share of any associated firm or joint venture R&D investment, are excluded. Furthermore, the accounts used are the consolidated group accounts of the ultimate parent company. Firms which are subsidiaries of any other company are not ranked separately. The specific data handling procedure incorporates the view that the crucial strategic decisions are taken by the central management of the firm and the degrees of freedom which remain for the "peripheral management" is rather limited to operational aspects (Penrose, 1959).

In order to examine the potential endogeneity between R&D and Export Intensity, for present purposes those firms that fail to report their exporting activity were excluded.[3] Thus, after cleaning the data, the dataset contains 498 firms and 780 observations in total. More specifically, 282 firms are reported for both of the two time periods available, while 68 firms are reported only for the year 2006, and 148

[3] At this point it should be noted that, based on the nature and reputation of the excluded firms, it is by no means groundless to assume that they are actively engaged in exporting activities but did not report the corresponding information.

firms are reported only for the year 2007. Given the particular data structure, it will be handled as pooled cross-section, according to which during each year a new random sample is taken from the relevant population.[4] The definitions of all the variables used in the estimations and the corresponding descriptive statistics are presented in Table 1 and Table 2, respectively.

In the present work, the relationship between R&D and Export Intensity will be tackled on a *global scale*. For that purpose, 'global industries' are constructed in the sense that hereafter, and with respect to the industry distribution, firms from all over the world will constitute a specific industry. In other words, the industry definition used in this paper is primarily based on the technological characteristics of the firms while significant intra-industry heterogeneity is allowed to be present, with respect to the country of origin, the geographical dispersion of the firm's activity, the organizational characteristics, and of course other fundamentals like size, profitability, etc. (Mansfield, 1964).

[4] This approach gives rise to independent, not identically distributed (i.n.i.d.) observations. However, it is important not to confuse a pooling of independent cross sections with a different data structure, panel data, where one follows the same group of individuals over time. In a pooling of cross sections over time, there is no replicability over time or, even if units appear in more than one time period, their recurrence is treated as coincidental and ignored (Wooldridge, 2002, pp. 128-9).

Table 1. Definition of Dependent and Explanatory Variables

Variable	Definition
	Dependent Variables
EXPINT	Revenues generated by Exports divided by firm's Sales
RDINT	Expenditures on R&D divided by firm's Sales
	Explanatory Variables
DEUR	A dummy variable = 1 if the firm's location is within Europe; = 0 otherwise
DNAM	A dummy variable = 1 if the firm's location is within North America, = 0 otherwise
DMAN1HT	A dummy variable = 1 if the firm is in the Manufacturing, High-tech sector; = 0 otherwise
DMAN2LT	A dummy variable = 1 if the firm is in the Manufacturing, Low-tech sector; = 0 otherwise
DCG	A dummy variable = 1 if the firm is in the Durable and Capital goods sector; = 0 otherwise
DSERV	A dummy variable = 1 if the firm is in the Services sector; = 0 otherwise
DICT1	A dummy variable = 1 if the firm is in the ICT sector; = 0 otherwise
$TD1$	A dummy variable = 1 if the time period is 2007; = 0 otherwise
RDT1	R&D intensity of the previous year, i.e. *t-1*
MSPREUR	Percentage of the firm's attained sales (exports) in Europe
MSPRNAM	Percentage of the firm's attained sales (exports) in North America
SALES00	Firm's size captured by its Sales
SALES002	Firm's size squared
HI	Sum of the squares of the market shares of each firm
LAB_PROD	Ratio of total revenues to number of employees, capturing labor productivity
PROFITAB	Firm's profitability index, Operating Profit divided by Market Capitalization (*OP/MarkCap*)
PROF_MAR	Another profitability ratio, Operating Profit divided by Sales $(OP/Sales)$
EXPINT1	Firm's total revenues from exporting in *t-1* divided by own Sales in *t-1*
EXPINT12	EXPINT squared

Table 2. Descriptive statistics of the variables used

Dependent Variables											
	Mean	**St Dev**	**Median**	**Min**	**Max**						
EXPINT	0.428	0.224	0.441	0.003	1.000						
RDINT	0.081	0.109	0.045	0.001	1.194						
Explanatory Variables											
	Mean	**St Dev**	**Median**	**Min**	**Max**		**Mean**	**St Dev**	**Median**	**Min**	**Max**
*TD*1	0.449	0.498	0.000	0.000	1.000	*HI*	0.072	0.066	0.050	0.026	1.000
DMAN1HT	0.512	0.500	1.000	0.000	1.000	*LAB _ PROD*	0.203	0.162	0.160	0.011	1.848
DMAN2LT	0.068	0.252	0.000	0.000	1.000	*MSPREUR*	0.374	0.256	0.340	0.000	0.991
DCG	0.117	0.321	0.000	0.000	1.000	*MSPRNAM*	0.267	0.198	0.220	0.000	0.997
DSERV	0.051	0.221	0.000	0.000	1.000	*EXPINT*1	0.423	0.225	0.430	0.004	1.000
DICT1	0.100	0.300	0.000	0.000	1.000	*EXPINT*12	0.229	0.209	0.185	0.000	1.000
DEUR	0.505	0.500	1.000	0.000	1.000	*PROFITAB*	0.062	0.243	0.077	-2.944	1.625
DNAM	0.232	0.422	0.000	0.000	1.000	*PROF_MAR*	0.090	0.172	0.093	-1.600	0.960
*SALES*00	0.063	0.132	0.018	0.000	1.429	*RDT1*	0.079	0.113	0.045	0.000	1.568
*SALES*002	0.021	0.123	0.000[5]	0.000	2.041						

5. Results and Discussion

In this section, we will present the empirical estimates of the econometric model as it was previously formulated. The structure of this section consists of two subsections. The first subsection is concerned with the empirical findings regarding the endogeneity between R&D and Export intensity. In the second subsection, FIML estimates of the export and R&D intensity determinants are presented and discussed. Nevertheless, before we proceed with discussing the results it is worth mentioning some issues and details related to the econometric estimations.

[5] Values are less than 0.0001

Firstly, on the basis of previous empirical findings regarding the determinants of R&D and export intensity, we have included a meaningful and informed set of explanatory variables among the economic and financial variables available. The number of control variables used in each equation indicated potential multicollinearity problems among the regressors, yet the Variance Inflation Factor (VIF) matrix shown in Table 3 suggests that no serious multicollinearity[6] problems seem to be present according to the corresponding values.

Secondly, in the context of the selection of variables for each equation an important issue needs to be addressed here. Despite the fact that the main focus for the potential existence of endogeneity lies in the first equation, where R&D intensity is a suspected endogenous explanatory variable, the relevant economic theory dictates that export intensity is also a determinant of R&D intensity (Bhattacharya and Bloch, 2002).

However, due to the recursiveness of the model it is not possible to introduce export intensity as a control variable in the equation where R&D intensity is the dependent variable.[7] Therefore, we decided to use *EXPINT1* as an Instrumental Variable (IV) in order to capture any variation with respect to R&D intensity. In the next subsection we will address this issue in more detail as it is crucial for

[6] In order to decide whether multicollinearity is not only present but also it really constitutes a problem there are two paths to choose from. The first path involves examining the VIF separately for each variable and if it takes large values, say over 10 (some authors say that the threshold is somewhere between 30 and 40) then there is a problem. The other path is to calculate the mean VIF of all the variables and if it is less than 10 then the model is free from multicollinearity. For further elaboration on this issue refer to Greene (2002, p. 57). Specifically in this case, the sole hint for multicollinearity problems lies between the variables *EXPINT1* and its square form *EXPINT1* 2. However, due to their economic and econometric significance and also taking into account the fact the overall VIF is equal to 3.291 we decided to include the square form into the model.

[7] After a personal contact with W. Greene he suggested that due to the recursiveness of the model it is not possible to include the left-hand variable of the first equation as a control variable in the second equation.

disentangling the relationship between these two activities and their underlying economic intuition.

Table 3. Variance Inflation Factors for the variables

Variance Inflator Factor (*EXPINT* Equation)		**Variance Inflator Factor (*RDINT* Equation)**	
Cons	0.000	*Cons*	0.000
DEUR	3.584	*DMAN1HT*	2.893
DNAM	2.271	*DMAN2LT*	1.731
*TD*1	1.015	*DCG*	2.165
SALES00	1.293	*DSERV*	1.950
SALES002	4.485	*DICT1*	1.895
LAB _ PROD	1.225	*DEUR*	1.753
RDINT	4.698	*DNAM*	1.987
MSPREUR	2.732	*TD*1	1.015
MSPRNAM	1.584	*SALES00*	1.233
PROFITAB	2.027	*SALES002*	4.399
PROF_MAR	2.200	*HI*	1.646
		EXPINT1	13.398
		EXPINT1 2	13.533
		RDT1	1.654
		PROF_MAR	1.408
Mean VIF	2.260	*Mean VIF*	3.291

In addition, the following econometric results, especially with respect to the relationship of the most important variables i.e. *EXPINT* and *RDINT*, remain consistent with previous estimations based on the cross-section dataset. Finally, in order to select the model with the best econometric properties among alternative models, a forward selection process was followed. This implies that some variables with no statistically significant coefficients have been included in the final model, as they are considered to be an important finding and because such an inclusion does not worsen the overall econometric performance of the model.

5.1 The EXPINT and RDINT relationship

Estimation results of the above two-equation model are presented in Table 4. At the bottom of the same table the estimated value and the corresponding asymptotic standard error of the parameter ψ, designed to test for the presence of endogeneity, are also displayed. According to the performed test, the hypothesis of weak exogeneity of *RDINT* with respect to *EXPINT* is not accepted. Thus, the degree of internationalization and the knowledge nature of the RDL firms are endogenously determined. Any further attempt to disentangle the relationship between them should explicitly acknowledge this fact.

Focusing on the *EXPINT* equation of the estimated model the most intriguing result is concerned with the positive and statistically significant coefficient of the *RDINT* variable. Recalling once more that the examined firms are global R&D leaders, this finding is important in many respects.

On the one hand it could be argued that the international orientation of the RDL firms is stimulated by the comparative advantage they possess from their knowledge base. On the other hand, one could argue that the RDL firms are "forced" to penetrate foreign markets – where the size of demand is theoretically approaching infinity – in order to exploit the steep economies of scale, thus reducing the unit cost associated with their R&D investment indivisibilities, or with the lumpiness of the technologies they produce. Of course, the above situations are not, in any case, mutually exclusive.

Table 4. Likelihood estimates from the simultaneous estimation of the *EXPINT* and *RDINT* equations

EXPINT Equation						*RDINT* Equation			
Variables	Coefficient Estimates	Marginal Effects	Variables	Coefficient Estimates	Marginal Effects	Variables	Coefficient Estimates	Variables	Coefficient Estimates
Cons	-0.356***	-	*LAB_PROD*	-0.017	-	*Cons*	0.051*	*DNAM*	0.021**
	(0,109)			(0.016)			(0.011)		(0.012)
DEUR	-0.0799	-	*RDINT*	7.395*	7.395*	*DMAN1HT*	-0.003***	*SALES00*	-0.250*
	(-0.085)			(0.823)	(0.823)		(0.001)		(0.073)
DNAM	-0.148	--	*MSPREUR*	-0.052***	-0.052***	*DMAN2LT*	-0.005***	*SALES002*	0.163***
	(0.095)			(0.021)	(0.021)		(0.002)		(0.067)
*TD*1	0.014	-	*MSPRNAM*	-0.062***	-0.062***	*DCG*	-0.003***	*HI*	0.033*
	(0.050)			(0.020)	(0.020)		(0.001)		(0.005)
SALES00	1.829***	1.829***	*PROF_MAR*	2.214*	2.214*	*DSERV*	-0.005***	*EXPINT*1	0.150*
	(0.605)	(0.605)		(0.262)	(0.262)		(0.002)		(0.018)
SALES002	-1.192***	-1.192***	*PROFITAB*	0.015	-	*DICT1*	-0.003***	*EXPINT*12	-0.030*
	(0.540)	(0.540)		(0.014)			(0.001)		(0.007)
						*TD*1	-0.001	*PROF_MAR*	-0.292*
							(0.006)		(0.007)
						DEUR	0.012	*RDT*1	0.023*
							(0.010)		(0.004)
ρ = -0.994		Log-Likelihood=3636.202		AIC = -9.247		$\sigma\left[\varepsilon_1 / \varepsilon_2\right] = 0.067^{**}$ (0.001)		ψ = -7.529*** (0.831)	

- Numbers in parentheses are the asymptotic standard errors
- One, two and three asterisks denote statistical significance at 1% and 5% and 10% respectively

Summarizing this part of discussion, it is evident that the two sides of the coin consist of the global and the knowledge producing firm, in the sense that the process of globalization is pumped by a strong and continuously expanding knowledge base.

Next we move towards the *RDINT* equation of the estimated model and take into account that the examined RDL firms are engaged in exporting activities. The question arising of whether their degree of internationalization exerts any impact on their R&D intensity is twofold: is there any impact exerted by exporting activities (i) on the knowledge creation process of RDL firms, and (ii) on the efficiency of the process itself? The *RDINT* equation of the estimated model provides some interesting insights into these issues. As previously mentioned, the variable *EXPINT*1 is included among the explanatory variables for both economic and econometric reasons. In particular, it captures the knowledge inflows of the RDL firms due to their previous exporting activities and constitutes an IV for the *EXPINT* variable.[8] It should be mentioned that the inclusion of such an IV in the *RDINT* equation raises a further question as to whether the endogenous relationship between export intensity and R&D intensity goes both ways. To the best of our knowledge, the available econometric techniques, including the one employed, do not allow us to deduce such a conclusion rigorously. However, we can recognize that there is some fertile soil for further investigation of this issue and that this relationship is multifaceted, and as yet far from disentangled.

In order to examine the endogenous relationship between *RDINT* and *EXPINT*1 more thoroughly, we have plotted these two variables against each other in

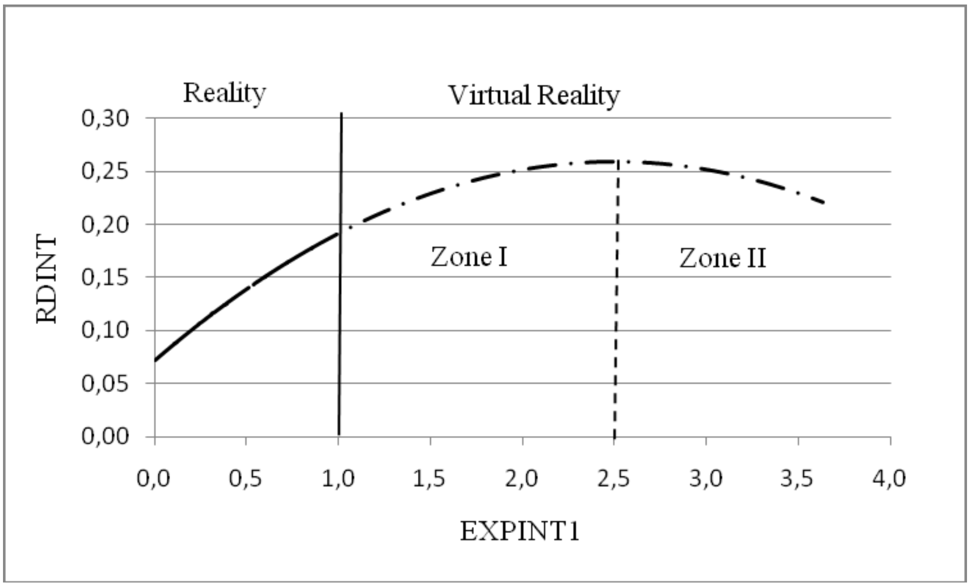
[8] The simple correlation coefficient between *EXPINT* and *EXPINT*1 variables equals to 0.987. This econometric strategy fulfils the necessary condition of the recursiveness of the estimated model.

Figure 1, holding all other statistically significant variables constant at their sample means according to the following equation:

$$RDINT_i = 0.071 + 0.150*\left(EXPINT1\right)_i - 0.031*\left(EXPINT1\right)_i^2 \quad (10)$$

It is quite evident that R&D and export intensity exhibit mixed patterns of interrelationship.

Figure 1. The knowledge creation process determined by the degree of RDL firms' internationalization

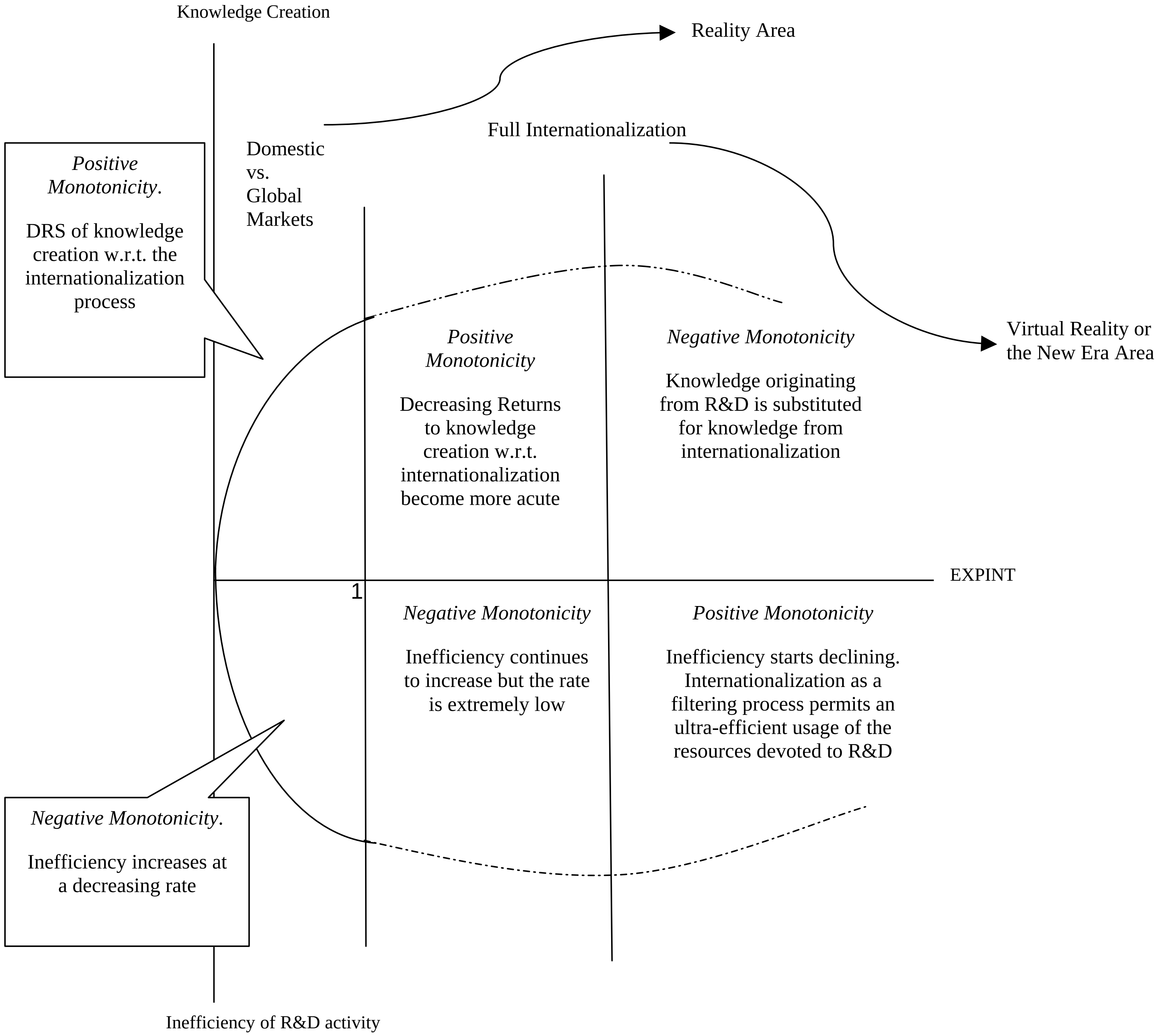


Two main areas are identified, namely the 'Reality' and the 'Virtual Reality' area. The area of 'Reality' essentially is where firms operate since it is not possible for any firm to have an export intensity index greater than one. In this area equation (10) is characterized by strict monotonicity and concavity. Economically speaking, R&D intensity, as a measure of RDL firms' knowledge creation, is ruled by decreasing returns to their international orientation. On the other hand, considering R&D intensity as a measure of the inefficiency of the resources devoted to R&D activities, it seems that higher export intensity leads to some decline in the efficiency of R&D activities but at a decreasing rate.

Moving to the '*Virtual Reality*' area where the export intensity index exceeds the value of one, we can identify two zones. In terms of Zone I, it may be regarded as an extrapolation of the 'reality area'. Positive monotonicity and concavity continue to be present. In this zone the RDL firm would continue to increase its R&D intensity due to its degree of internationalization though at a decreasing rate. Inefficiency of the R&D activities continues to be exacerbated but the rate is decreasing even more than the corresponding trend of the reality area.

Zone II depicts the imaginary situation of a new era arising from a drastic structural change, where the RDL firm has reached the point of maximum inefficiency of the resources devoted to augmenting its stock of knowledge. For values of export intensity greater than 1.0, which of course is unattainable, the RDL firm achieves such an exploitation of the knowledge flows derived from exporting activities that it permits the reduction of the intensity of its own R&D activities. In terms of orthodox economic theory one could argue that Zone II of the 'Virtual Reality' area represents a 'technology' where the knowledge inflows from exporting, compared to the corresponding level of R&D activities, are close substitutes. Some kinds of filtering mechanisms of the R&D activities through knowledge acquired by the RDL firm's internationalization are in operation here. A more thorough representation of our theoretical argumentation is presented in Figure 2 in which we have eliminated most of the quantitative and technical detail.

Figure 2. Illustration of the interrelationship between R&D and export intensity of the RDL firms



It would be quite interesting to establish the conditions under which the RDL firms would be able to overcome the tough reality area and even more reach the turning point of Zone II. The paper at hand allows us only to speculate on these conditions. In this direction, it is worth mentioning that the following may each play some role in formulating the underlying conditions of this relationship: (i) the organizational structures of the examined firms; (ii) the specific features of the economic environment (level of competition, trade flows, asymmetries of information, international financial systems, etc.); (iii) the specific idiosyncrasies of the technology development processes; and (iv) the alternative possible ways in which RDL firms secure access to knowledge.

5.2 The remaining determinants of EXPINT and RDINT

Besides the main relationship between R&D and export intensity, which was analytically presented and discussed above, each set of control variables for each equation of the system will be the interest of this section. In the two equations, we have included 6 common variables for both equations and 16 that are present only in one of the two equations. Specifically, in our case, where the R&D leader firms are the centre of interest, the econometric results reveal some interesting findings.

Regarding the common set of variables, two out of the 6 variables are found not to be statistically significant, i.e. the dummy variable of the country of origin $(DEUR)$ and the time dummy $(TD1)$. At this point it can be noted that the variable *DNAM* is positive and statistically significant, but only when used as a determinant of R&D intensity. This finding confirms the general conception that Europe is generally not the centre of technological advance, in contrast to the USA and Japan.[9] The statistical non-significance of *TDI* is not so surprising, since it involves only two time

9 For that reason, in the Lisbon (2000) convention, the European Union revealed its intentions to become "the most competitive and dynamic knowledge based economy in the world, capable of sustainable economic growth with more and better jobs and greater social cohesion." See "Presidency Conclusions", Lisbon European Council, 23 and 24 March 2000, Press Release Library, European Commission.

periods and any existing variation due to time may be absorbed by the other control variables.

In terms of the *SALES* $(SALES00, SALES002)$ variables, they are included in both equations as a proxy for firm size since there is a vast literature which has attributed it considerable importance in explaining the variation of both R&D and export intensity. In this light, it is interesting to examine the effects of firm size on the export intensity, taking into account that significant quadratic effects are found. More specifically, the estimated parameters of *SALES00* and *SALES002* are statistically significant at the 0.05 level, but the estimated coefficient on *SALES00* exerts a positive influence while that on *SALES002* has a negative influence on export intensity. Therefore, it is evident that this relationship too, between size and export intensity, is nonlinear: when firm size is relatively small, exporting activities have a positive effect but only up to a certain threshold, after which exporting becomes negatively correlated with firm size.

This finding is in accordance with the findings of several authors (Schlegelmilch and Crook, 1988; Kumar and Siddhartan, 1994; Wagner, 1995; Wakelin, 1998). Wagner (1995) argued that firm size advantages are present up to a certain threshold due to coordination costs and bureaucratic issues, while Schlegelmilch and Crook (1988) argued that this nonlinearity is due to the fact that above a certain size large firms find it more efficient to proceed to FDI rather than exporting. These explanations might not be contradictory but complementary. More specifically, it can be argued that these firms have actually reached a certain size which offers them the advantage of having alternative choices such as FDI or intra-firm trade but, on the other hand, the decisions need to be made have become more complicated.

From the *RDINT* equation, the econometric results regarding the *SALES00* and *SALES002* variables are also quite interesting. While both are statistically significant, *SALES00* is negatively correlated with R&D intensity while *SALES002* is positively correlated. In this case, regarding the firms which are considered R&D leaders, it seems that being big may be an advantage over being small with respect to R&D intensity, in the sense that the firm must devote a considerable amount from its available resources in order to be engaged in R&D activities.

The *PROF_MAR* variable which constitutes an indicator of the market power[10] a firm may possess and exert (Sullivan, 1985), when included in the *EXPINT* equation, is positive and statistically significant at the 0.05 level. According to this empirical finding the argument that international trade contributes to the increase of the welfare due to international competitive forces does not seem to apply in this case. Market power having a positive influence on export intensity suggests that the competition on a global scale or outside the domestic market does not resemble the perfect competition model but on the contrary the oligopolistic model is more suitable to explain firms' export behavior. Furthermore, the same variable exerts negative and statistically significant impact when included as a determinant of R&D intensity. It may be argued that, for the R&D leader firms, the possession of market power at the firm level deflects the intensity of R&D, implying that these firms need to feel the 'scent of competition' in order to intensify their R&D expenditures.

As far as the rest of the determinants of the *EXPINT* equation are concerned, two variables, *LAB_PROD* and *PROFITAB*, were inserted in order to test for the self-

[10] It should be pointed out that this market power refers to both a domestic and a global scale.

selection hypothesis (Wagner and Strasse, 2007) and according to the econometric results this hypothesis cannot be accepted since neither the labor productivity index nor the RDL firms' profitability is statistically significant.

On the other hand, two other variables were used in order to capture any effects of knowledge flows through exporting (UNCTAD, 2005). More specifically, it is argued that exporting firms will be in a position to exploit the potentials that the destination country offers them. In this case, the variables *MSPREUR* and *MSPRNAM* represent the firms' total exports to the respective region. The results indicate that both these variables are negative and statistically significant. The negative sign of the two variables can only be interpreted in relation to the omitted variable, *MSPRROW*. If one considers the hypothesis that exporting activities operate as a channel for technology transfer, it seems that exporting to the European and North American regions in contrast to exporting to the Rest of the World (RoW) does not have a positive impact on export intensity. This would imply that the most prominent region for acquiring or transferring knowledge and technology is the RoW region, which includes the fast developing countries of Japan, Korea, etc. Alternatively the connection here may go through the demand side, in that firms exporting to European or North American destinations may face extra competition from indigenous firms within those regions.

Turning to the determinants used only in the R&D intensity equation, the results of the estimated coefficients show that all the industry dummies are statistically significant and negatively affect the R&D intensity. This is not a surprising result if one considers that the omitted industry dummy variable, which is used as a reference point, is the ICT2 variable (i.e. the hardware branch of ICTs).

Moreover, since all industry dummies are related to the base case this negative relationship may simply indicate values that are less than that of the base industry.

The *RDT1* variable depicts the R&D intensity of the year *t-1*. It takes a positive sign and is statistically significant at 0.01 level. This finding comes to support the notion of path dependency. R&D projects present certain characteristics and among those are that they are complex and time-consuming (Hobday, 2000). In other words, once an RDL firm has been engaged in an R&D project in the past, it is more than likely that it will continue working on the same or similar R&D project in the future, *ceteris paribus*.

Finally, the industrial concentration was measured with the Herfindahl index (HI) in order to test for the Schumpeterian hypothesis concerned with the effect of industry concentration with respect to innovation. Results of the estimated coefficient indicate that in this case the Schumpeterian hypothesis is confirmed as it is positive and statistically significant. This empirical finding implies that the more concentrated a global industry is the more positive impact it will have for the level of R&D intensity. This may be due to the fact that the industrial concentration ensures appropriability conditions, minimizing free-rider problems and other kinds of uncertainties related to R&D activities. Overall, we could argue that innovation is probably the most important means for the firm to preserve its monopoly power in a global market.

6. Conclusions

This paper is concerned with shedding some light on some of the more obscure aspects of the relationship between the internationalization process, captured by export intensity, and the knowledge creation process, captured here by R&D intensity, as regards the R&D leader firms on a global scale. In this respect, we have estimated a simultaneous two-equation model where one right-hand variable is censored. Empirical results reveal that export intensity is endogenously determined by R&D intensity and in addition, the knowledge creation process affects positively the degree of internationalization. It is interesting to notice that this finding, along with the remaining determinants, have been interpreted based on the fact that these firms are not restricted to their domestic markets but on the contrary they form and compete in a global market in terms of technology creation and knowledge acquisition. To this end, we have argued that the levels of internationalization and knowledge creation serve the purpose of augmenting the RDL firms' pool of knowledge, among other impacts.

The issues become more complicated when in the aforementioned endogeneity framework we incorporate the impact of the degree of internationalization on the knowledge creation process. A nonlinear relationship has been identified between these two activities and decreasing returns with respect to knowledge creation. More specifically, two areas namely 'reality' and 'virtual reality' arise; however firms are doomed to operate in the 'reality' area. The 'virtual reality' area can be separated into two zones and what is most interesting is that under extraordinary circumstances, translates into drastic structural changes, RDL firms would be able to substitute the knowledge flows from R&D activities with those flows to be derived from exporting activities, improving the efficiency of their knowledge creation at the same time.

It should be noted that these empirical findings gain meaning only in a context of a global market. Therefore, these firms, by being R&D leaders, have a special trait that the majority of firms in this global context do not have. Further research is needed to investigate the endogeneity between R&D and export intensity in different contexts, i.e. differences in industry, country, technology, time specificity, and/or with reference to other types of firms. It would also be useful to pursue the decomposition of these knowledge flows with respect to the internal resources devoted to creating knowledge and also with respect to the decreasing returns to knowledge creation by exporting activities. Last but not least, further research is needed regarding the organizational features of RDL firms which are responsible for the identified relationships between the examined knowledge flows.

References

Aghion, P. and Howitt, P. (1998). *Endogenous Growth Theory*, Cambridge MA: MIT Press.

Bhattacharya, M. and Bloch, H. (2004). Determinants of Innovation, *Small Business Economics*, 22/2, 155-162.

Bleaney, M. and Wakelin, K. (2002). Efficiency Innovation and Exports, *Oxford Bulletin of Economics and Statistics*, 64/1, 3-15.

Clerides, S., Lach, S. and Tybout, J. (1998). Is Learning by Exporting Important? Micro-dynamic Evidence from Colombia, Mexico, and Morocco, *Quarterly Journal of Economics*, 113/3, 903–47.

Cohen, W. and Levin, R, (1989). Empirical Studies of Innovation and Market Structure. In Schmalensee, R. and Willig, R. (eds), *Handbook of Industrial Organization*, Elsevier Science Publishers, ch. 18.

Cohen, W. and Levinthal, D. (1989). Innovation and Learning: the Two Faces of R&D, *The Economic Journal*, 99/September, 569-96.

Dollar, D. (1986). Technological Innovation, Capital Mobility, and the Product Cycle in North-South Trade, *American Economic Review*, 76/1, 177-190.

Evenson, R. and Westphal, L. (1995). Technological Change and Technology Strategy. In Behrman, J. and Srinivasan, T.N. (eds), *Handbook of Development Economics, 3*, Amsterdam: North-Holland, 2209–98.

Färe, R., Grosskopf, S., Norris, M. and Zhang, Z. (1994). Productivity Growth, Technical Progress, and Efficiency Change in Industrialized Countries, *American Economic Review*, 84/ 1, 66-83.

Farrell, M. J. (1957). The Measurement of Productive Efficiency, *Journal of the Royal Statistical Society, Series A (General)*, 120/3,253-290

Gkypali, A. (2008). *The R&D and Export Intensity: a Puzzling Relationship*. M.Sc Dissertation, SPRU, UK.

Grant, R.M. (1996). Toward a Knowledge-based Theory of the Firm, *Strategic Management Journal*, 17 (Winter Special Issue), 109–122.

Greene, W.H. (2002). *Econometric Analysis.* New Jersey: Prentice Hall.

Greene, W.H. (2007). *Limdep v.9. Econometric Modeling Guide*, 2, E27-58. New York: Econometric Software.

Grossman, G. and Helpman, E. (1989). Product Development and International Trade, *Journal of Political Economy*, 97/6, 1261–83.

Grossman, G. and Helpman, E. (1990). Comparative Advantage and Long-run Growth, *American Economic Review*, 80/4, 796–815.

Grossman, G. and Helpman, E. (1991). Endogenous Product Cycles, *Economic Journal*, 101/September, 1214–29.

Harris, R. and Li, Q. (2008), Exporting, R&D, and Absorptive Capacity in UK Establishments, *Oxford Economic Papers*, 69/1, 74-103.

Heckman, J. (1978). Dummy Endogenous Variables in a Simultaneous Equation System, *Econometrica*, 46/6, 931-959.

Hobday, M. (2000). The Project-based Organisation: an Ideal Form for Managing Complex Products and Systems? *Research Policy*, 29/7-8, 871-893.

Johanson J. and Vahlne, J.E. (1977). The Internationalization Process of the Firm: A Model of Knowledge Development and Increasing Foreign Market Commitments, *Journal of International Business Studies*, 8/2, 23-32.

Keesing, D. (1967). The Impact of Research and Development on United States Trade, *Journal of Political Economy*, 75/1, 38–48.

Kogut, B. and Zander, U. (1992). Knowledge of the Firm, Combinative Capabilities, and the Replication of Technology, *Organization Science*, 3/3, 383-397.

Krugman, P. (1979). A Model of Innovation, Technology Transfer, and the World Distribution of Income, *Journal of Political Economy*, 87/2, 253-66

Kumar, N. and Siddharthan N. (1994). Technology, Firm Size and Export Behaviour in Developing Countries: The Case of Indian Enterprises, *Journal of Development Studies*, 31/2, 289-309.

Lachenmaier, S. and Woessmann, L. (2006). Does Innovation Cause Exports? Evidence from Exogenous Innovation Impulses and Obstacles using German Micro Data, *Oxford Economic Papers*, 58/2, 317–350.

Lall, S. and Kumar, R. (1981). Firm-level Export Performance in an Inward-looking Economy: The Indian Engineering Industry, World Development, 9/5, 453-463.

Lefebvre, E., Bourgault, M. and Lefebvre, L. (1998). R&D Related Capabilities as Determinants of Export Performance, *Small Business Economics*, 10/4, 365-377.

Mansfield, E. (1964). Industrial Research and Development Expenditures: Determinants, Prospects, and Relation of Size of Firm and Inventive Output, *Journal of Political Economy*, 72/4, 319-340.

Nonaka, I. (1994). A Dynamic Theory of Organizational Knowledge Creation, *Organization Science*, 5/1, 14-37.

Pedersen, T., Petersen, B. and Sharma, D.D. (2003). Knowledge Transfer Performance of Multinational Companies, *Management International Review*, 43/3, 69-90.

Penrose, E.T. (1959). *The Theory of the Growth of the Firm*, New York: Wiley.

Pla-Barber, J. and Alegre, J. (2007). Analysing the Link between Export Intensity, Innovation and Firm Size in a Science-based Industry', *International Business Review*, 16/3, 275-293.

Posner, M. (1961). International Trade and Technical Change, *Oxford Economic Papers*, 13, 323-34.

Saggi, K. (2002). Trade, Foreign Direct Investment and International Technology Transfer: a Survey, *World Bank Observer*, 17/2, 191-235.

Schlegelmilch, B. and Crook, J. (1988). Firm-Level Determinants of Export Intensity, *Managerial and Decision Economics*, 9/4, 291-300.

Schumpeter, J. (1942). *Capitalism, Socialism and Democracy*, New York: Harper and Row.

Segerstrom, P., Anant, T. and Dinopoulos, E. (1990). A Schumpeterian Model of the Product Life Cycle, *American Economic Review*, 80/5, 1077–91.

Smith, R. and Blundell, R. (1986). An Endogeneity Test for a Simultaneous Equation Tobit Model with an Application to Labour Supply, *Econometrica*, 54/3, 679-685.

Smith, V., Madsen, E. and Dilling-Hansen, M. (2002), Do R&D Investments Affect Export Performance? *WP 2002/4*, The Danish Institute for Studies in Research and Research Policy, Aarhus.

Spender, J. (1996). Making Knowledge the Basis of a Dynamic Theory of the Firm, *Strategic Management Journal*, 17 (Winter Special Issue), 45-62.

Sullivan, D. (1985). Testing Hypotheses about Firm **B**ehavior in the Cigarette Industry, *Journal Political Economy*, 93/3, 586–598.

UNCTAD (2005). *Determinants of Export Performance*, Developing Countries in International Trade, Trade and Development Index.

van Dijk, M. (2002). The Determinants of Export Performance in Developing Countries: The Case of Indonesian Manufacturing, *Working paper for Eindhoven Centre for Innovation Studies (Ecis)*.

Vernon, R. (1966). International Investment and International Trade in the Product Cycle, *Quarterly Journal of Economics*, 80/2, 246-263.

Wagner, J. (1995). Exports, Firm Size, and Firm Dynamics, *Small Business Economics*, 7/1, 29-39.

Wagner, J. (2001). A Note on the Firm Size Export Relationship, *Small Business Economics*, 17/4, 229-237.

Wagner, J. and Strasse, K. (2007). Exports and Productivity: A Survey of the Evidence from Firm-Level Data, *World Economy*, 30/1, 60-82.

Wakelin, K. (1998). Innovation and Export Behaviour at the Firm Level, *Research Policy*, 26/7-8, 829-841.

Wilmore, L. (1992). Transnationals and Foreign Trade: Evidence from Brazil, *Journal of Development Studies*, 28, 314-335.

Wooldridge, J.M. (2002). *Econometric Analysis of Cross Section and Panel Data*. Cambridge MA: MIT Press.

Young, A. (1991). Learning by Doing and the Dynamic Effects of International Trade, *Quarterly Journal of Economics*, 106/2, 369-406.