On the Link Between R&D, Innovation and Productivity: Panel Evidence for Dutch and French Manufacturing

Wladimir Raymond[*], Jacques Mairesse[†], Pierre Mohnen[‡] and Franz Palm[§]

(Preliminary draft)

February 7, 2010

Abstract

This paper studies the relationship between R&D, innovation and productivity in Dutch and French manufacturing using an unbalanced panel of enterprise data from three waves of the Community Innovation Survey pertaining to the period 1994-2004. We estimate by FIML a simultaneous-equation model of R&D, innovation and productivity accounting for individual effects and correcting for sample selection bias. Four-year lagged R&D affects positively and significantly innovation which itself is positively related to productivity. Individual effects play an important role in the relationship and selectivity bias must be accounted for. Significant differences across countries are observed.

[*]University of Maastricht, P.O. Box 616, 6200 MD Maastricht, The Netherlands; Room A4.27; Tel.: (+31) 43 388 4962; Fax: (+31) 43 388 4874; w.raymond@maastrichtuniversity.nl

[†]CREST-INSEE, NBER and Maastricht University; mairesse@ensae.fr

[‡]Cirano and Maastricht University, p.mohnen@maastrichtuniversity.nl

[§]CESifo fellow and Maastricht University, f.palm@maastrichtuniversity.nl

1 Introduction

In the early sixties, several empirical studies adopt the aggregate production function framework of Solow (1957) to explain economic growth in the US, and find that most of the growth is attributed to the residual factor labeled as 'technical change'.[1] Griliches (1963) feels dissatisfied with this approach and argues that the contribution of the residuals to economic growth can be reduced by, among others, a better specification of the aggregate production function. In his quest for such a specification, Griliches (1979) augments the production function with a 'knowledge capital' input which he proxies by the stock of R&D capital. Following the latter approach, many studies investigate the return of output (growth) to R&D (see Griliches and Mairesse, 1984). Another strand of literature, somewhat related to the previous one, is about the relationship between innovation input (e.g. R&D) and innovation output (e.g. sales of new products) also known as knowledge production function (henceforth KPF). The concept of KPF, although not referred to as such, dates back in the mid-sixties (e.g. Scherer, 1965) and is popularized by among others Pakes and Griliches (1980)

The Crépon-Duguet-Mairesse (1998) (in short CDM) model unifies the two above-mentioned approaches into a single framework where the determinants of innovation and their effects on economic performance are studied. These two questions probably form the most important research topic dealing with innovation survey data. Many studies (around 50 to 100) on data from many countries estimate variants of the standard CDM model. The popularity of this model is due to its ability to answer the two questions in a single framework, and to account for the endogeneity of innovation input and innovation output, and the censoring/truncated feature of the data derived from the innovation surveys based on the Oslo Manual (OECD, 2005). So far, however, the CDM model is essentially estimated on cross-sectional data to explain the level or the growth of labor productivity. Only a few studies are based on panel data, but partially or not at all accounting for two characteristics inherent to the innovation process, namely individual effects (e.g. unobserved ability to innovate) and dynamics (e.g. lag effect of R&D).

We derive a model where product innovation is explained by four year lagged R&D intensity, and where the latent variable of product innovation intensity explains labor productivity. The model is composed of three equations, namely a selection equation of product innovators, an equation for the intensity of product innovation and a labor productivity equation. We consider random individual effects in each equation, and compare two specifications of the distribution of

[1] The study of Solow is the most widely cited when it comes to assess the large contribution of the residuals to economic growth. However, other economists, e.g. Abramovitz (1956), expressed that idea prior to Solow (see Griliches, 1994).

these random effects. The model is estimated by full information maximum likelihood (FIML) using two- and three-step Gauss-Hermite quadrature to handle technically the integration with respect to the bivariate and trivariate distributions of the individual effects. The determinants and the effects of product innovation are compared for an unbalanced panel of enterprise data for the manufacturing sector of France and The Netherlands using three waves of the Community Innovation Surveys, CIS 2, CIS 3 and CIS 4. We find a positive and significant of R&D on the share of innovative sales, even after four years, and a positive and significant effect of the latter on sales per capita. Individual effects play a significant role in all three equations and selectivity bias must be accounted for. Significant differences across countries are observed.

The remainder of the paper is organized as follows. In Section 2 we review the empirical literature using the CDM model. In Section 3 we describe the data and some descriptive statistics on the main variables of the model. The econometric specification with its two variants is presented in Section 4, followed by the results in Section 5. We conclude in Section 6.

2 Literature

Table 1 surveys the empirical literature on the interrelationship at the enterprise level between innovation input, innovation output and economic performance. The surveyed studies are variants of the CDM model applied to not only OECD countries (e.g. France, Germany, or Italy) but also to non-OECD countries (e.g. China, Chile).

The main message of the table is that the majority of those studies use cross-sectional data, hence failing to account (at all or properly) for two inherent features of the innovation process, namely dynamics and unobserved heterogeneity. With panel data, more realistic and flexible variants of the CDM model can be specified (cf. Figure 1 of Appendix A). Three types of dynamics can be modeled and unobserved heterogeneity can be handled by individual effects. First, due to the risky characteristic of innovation activities, innovation input may translate into innovation output only after a certain time (see e.g. Pakes and Griliches, 1980). All the studies of Table 1 assume an instantaneous effect of innovation input on innovation output, except Jefferson et al. (2006) who estimate a one-year lag effect of innovation input on innovation output. Another type of dynamic feature of the innovation process is its persistence that may occur in innovation input because of *sunk costs*, or in innovation output because of, among others, financial constraints or "success breeds success". Very few studies account for this type of dynamics in the innovation process. van Leeuwen (2002) and Jefferson et al. (2006) model the persistence of innovation input, and these authors and Parisi et al. (2006) model the persistence of innovation output. However,

Table 1: Surveyed literature on the interrelationship between innovation input, innovation output and economic performance[†]

Study	Country (Period)	Sector	Data source	Dependent variables: Innov. input	Innov. output	Performance	Dynamics	Individual effects
				Cross-sectional data				
Crépon et al., 1998	France (86-90)	Manuf.	CIS 0, EPAT	R&D/EMP	Patents/EMP (IS/Sales)[‡]	VA/EMP	no	no
Klomp and van Leeuwen, 2001	Netherlands (94-96)	Manuf. others	CIS 2, PS	TIE/Sales	IS/Sales	ΔSales ΔEMP	no	no
Lööf and Heshmati, 2002	Sweden (94-96; 95-97)	Manuf.	NUTEK's survey CIS 2, BR	(R&D/Sales)[‡] TIE/EMP	IS/EMP ISM/EMP IS/Sales	VA/EMP, Δ(VA/EMP) Δ(Sales/EMP)	yes[†‡]	no
Benavente, 2006	Chile (95-98)	Manuf.	ENIA, EIT	R&D/EMP	(IS/Sales)[‡]	VA/EMP	no	no
Griffith et al., 2006	France, Spain Germany, UK (98-00)	Manuf.	CIS 3	R&D/EMP	INPDT INPCS	Sales/EMP	no	no
van Leeuwen and Klomp, 2006	Netherlands (94-96)	Manuf.	CIS 2, PS	R&D/Sales TIE/Sales	IS/Sales	VA/EMP, Δ(VA/EMP) Δ(Sales/EMP)	no	no
Lööf and Heshmati, 2006	Sweden (96-98)	Manuf. serv.	CIS 2.5, BR	TIE/EMP	IS/EMP ISM/EMP	Profit/EMP, Δ(Profit/EMP) Sales/EMP, Δ(Sales/EMP) (VA/EMP), Δ(VA/EMP) Sales margin, ΔEMP	yes[†‡]	no
				Panel data				
van Leeuwen, 2002	Netherlands (94-98)	Manuf.	CIS 2, 2.5, PS	R&D/Sales	IS/Sales	Δ(Sales/EMP)	yes	no
Jefferson et al., 2006	China (97-99)	Manuf.	3 waves of NBS's innov. survey	TIE/Sales	IS/Sales	Profit/Sales, GrO/EMP	yes	no
Parisi et al., 2006	Italy (92-97)	Manuf.	6^{th} and 7^{th} waves of MCC's IIM	R&D/TK	INPDT INPCS	Δ(GrO/EMP), ΔTFP	yes (no)*	no (yes)*

[†] *Notes*: CIS=Community Innovation Survey, EPAT=European Patent Office, EMP=employment, IS=innovative sales, VA=value-added, TIE=total innovation expenditures, PS=production survey, Δ=annual growth, BR=business register, NUTEK=Swedish National Board for Industrial Technical Development, (E)NIA/IT= *'(Encuesta) Nacional Industrial Anual/de Inovacion Tecnologica'*, INPDT (INPCS)=1 if product (process) innovator, ISM=innov. sales, products new to the market, NBS=National Bureau of Statistics, MCC=*'Mediocredito Centrale'*, IIM=*'Indagine sulle Imprese Manifatturiere'*, GrO=gross output, TK=total capital, i.e. physical capital + R&D capital. [‡]given as interval data. [†‡]$\Delta(VA/EMP)_{-1}$ is included in the Δ(VA/EMP) equation. *This study accounts for either individual effects or dynamics, but not for both.

Table 1: Surveyed literature on the interrelationship between innovation input, innovation output and economic performance (cont.)†

Study	Methodology	Result		
		Innov. input	Innov. output	Performance
			Cross-sectional data	
Crépon et al., 1998	ALS; type 2 tobit, NegBin, Oprobit, linear regression	EMP^{ns}, MKT^{+}	EMP^{ns}, $R\&D^{+}$ (for both measures of innovation output)	EMP^{ns}, $(IS/Sales)^{+}$ $(Patents/EMP)^{+}$
Klomp and van Leeuwen, 2001	ML, type 2 tobit; OLS, FIML with/no Mill's λ, linear reg.	$(\Delta Sales)^{+}$, $Sales^{-}$ EMP^{ni}, MKT^{ni}	$(TIE/Sales)^{+}$, $Sales^{ns}$, $(\Delta Sales)^{ns}$	$(IS/Sales)^{+}$ for ΔSales $(IS/Sales)^{-}$ for ΔEMP
Lööf and Heshmati, 2002	2-step Heckman, type 2 tobit; OLS with Mill's λ, linear regression	$EMP^{ni,ns}$, $MKT^{ni,ni}$, $\Delta(Sales/EMP)^{+,ni}$ for R&D/Sales and TIE/EMP resp.	$(Sales/EMP)^{ni,ni,+}$, $(R\&D/Sales)^{ni,ni,+}$, $\Delta(Sales/EMP)^{ni,ni,+}$, $(VA/EMP)^{ns,+,ni}$, $\Delta(VA/EMP)^{-,ni,ni}$, $(TIE/EMP)^{+,ns,ni}$, $EMP^{ns,ns,+}$ for the 3 respective measures of innov. output	$(IS/Sales)^{+}$ for $\Delta(Sales/EMP)$; $(IS/EMP)^{+,+}$ for VA/EMP and ΔVA/EMP; $(ISM/EMP)^{+}$ for ΔVA/EMP; $EMP^{ns,ns,-}$ for the 3 respective perf. measures
Benavente, 2006	ALS; type 2 tobit, NegBin, Oprobit, linear regression	EMP^{ns}, MKT^{ns}	EMP^{ns}, $R\&D^{ns}$	EMP^{+}, $(IS/Sales)^{ns}$
Griffith et al., 2006	ML, type 2 tobit and probit; OLS, linear regression	EMP^{ni}, MKT^{ni}	EMP^{+} for INPCS (all 4 countries), EMP^{+} for INPDT (only in Spain), $(R\&D/EMP)^{+}$	EMP^{+}, $INPDT^{+}$ (not in Germany), $INPCS^{+}$ (only in France),
van Leeuwen and Klomp, 2006	ML, type 2 tobit; OLS with no λ, FIML with λ, 3SLS with/without λ, linear reg.	$EMP^{ns,-}$, $MKT^{ns,ns}$, $(\Delta Sales)^{ns,+}$ for both resp. input measures	$(R\&D/Sales)^{+}$, $(TIE/Sales)^{+}$; $EMP^{-,ns}$ when using respectively R&D and TIE as input	$(IS/Sales)^{ns,ns,+}$, $EMP^{+,+,+}$ for the 3 respective performance measures
Lööf and Heshmati, 2006	ML, type 2 tobit; OLS and 3SLS with/without Mill's λ	EMP^{-} for both manuf. and serv.	$(TIE/EMP)^{+}$, EMP^{+} (only for serv., *ns* for manufacturing)	$(IS/EMP)^{+}$ all measures but Sales marg., Δprofit and ΔEMP (manuf.); Sales marg., ΔSales (serv.); (ISM/ EMP)$^{+}$ for Sales, ΔSales (manuf.) VA, ΔVA, Profit (manuf.), ΔProfit & ΔEMP (serv.); EMP^{ns} for VA/EMP
			Panel data	
van Leeuwen, 2002	ML, type 2 tobit; FIML with Mill's λ, linear regression	EMP^{-}, $(R\&D/Sales)^{+}_{-1}$ MKT^{+}, $(IS/Sales)^{ns}_{-1}$	EMP^{ns}, $(IS/Sales)^{+}_{-1}$; $(R\&D/Sales)^{+}$ only if no dynamics, or else *ns*	EMP^{+} and $(IS/Sales)^{+}$ only if no dynamics in innov., or else *ns*
Jefferson et al., 2006	OLS, IV with no selection bias correction; linear reg.	TIE^{+}_{97}, $(Profit/Sales)^{ns}_{97}$, $Sales^{ns}_{97}$ in TIE_{98}	TIE^{+}_{98} in $(IS/Sales)_{99}$	$(IS/Sales)^{+}_{99}$ and EMP^{+}_{99} in both $(Profit/Sales)_{99}$ and $(GrO/EMP)_{99}$
Parisi et al., 2006	ML, logit (pooled, FE, RE); OLS, IV, linear reg. with no selection bias correction	not modeled	$(R\&D/TK)^{+,ns}$ (+ only in INPDT with no dynamics, *ns* in INPCS in all cases); $INPDT^{+}_{-1}$ in INPDT, $INPCS^{ns}_{-1}$ in INPCS	$INPDT^{+}$ only for ΔTFP; $INPCS^{+}$, $(INPDT\text{-or-}INPCS)^{+}$ for both measures

† *Notes*: Superscripts +, −, *ns*=positive, negative, non-significant effects, *ni*=not included; (A), (O), 3SLS=(asymptotic), (ordinary), 3-stage least squares, (FI)ML=(full-information) maximum likelihood, NegBin=negative binomial, Oprobit=ordered probit, FE and RE=fixed- and random-effects, IV=instrumental variables.

as it is very well documented in the panel data econometrics literature, genuine persistence can be ascertained only after accounting for unobserved persistent effects (e.g. individual effects) in the errors, which none of the listed studies has been able to achieve.[2] We may also model the persistence of economic performance for the same reasons as for innovation output but ensuring that individual effects are taken into account. Lööf and Heshmati (2002, 2006) estimate the effect of past productivity growth, as measured by value added over employment, on current productivity growth but are not able to control for individual effects. They find an insignificant effect, both economically and statistically. Finally, a feedback effect of past economic performance on current innovation activities may be modeled, also reflecting the financial constraints hypothesis. Indeed, because of information asymmetry, seeking external sources of funds may be more expensive than relying on retained earnings, which makes past profits condition current innovation activities (see e.g. Bhattacharya and Ritter, 1983). Some of the listed studies try to model this 'feedback' effect, but using cross-sectional data (e.g. Klomp and van Leeuwen, 2001). This renders the CDM model more complicated than necessary, as no causal relationship can be inferred from that effect. The only exception is Jefferson et al. (2006) who estimate an elasticity of total innovation expenditures in 1998 with respect to profit in 1997 of 0.03% but insignificant at the 5% level.

3 Data

The data used in the analysis stem from three waves of the Dutch and French CIS pertaining to the manufacturing sector for the periods 1994-1996 (CIS 2), 1998-2000 (CIS 3) and 2002-2004 (CIS 4). The Dutch CIS data are merged with data from the Production Survey (PS) that provides information regarding employment and sales of the firm, and both surveys are carried out by the *'Centraal Bureau voor de Statistiek'* (CBS). The French CIS data is collected by the *'Service des Statistiques Industrielles'* (SESSI) for the manufacturing sector (excluding the food industry). For each CIS, the merged PS variables pertain to the last year of the three-year period. We consider enterprises that belong to the manufacturing sector excluding the food industry, i.e. NACE 17.1-37.2, with at least ten employees and positive sales at the end of each period covered by the innovation survey. Only enterprises with a share of total R&D expenditures (intramural + extramural) in total sales smaller than or equal to 50% are included in the analysis.

The Dutch CIS and PS, and the French CIS data are collected at the enterprise level. A combination of a census and a stratified random sampling is used for each wave of both surveys. A

[2] Parisi et al. (2006) consider two specifications of innovation output, namely a dynamic specification where persistence is modeled but individual effects are ignored and a static specification where individual effects are accounted for.

census is used for the population of Dutch enterprises with at least 50 employees, and a stratified random sampling is used for enterprises with less than 50 employees, where the stratum variables are the economic activity and the number of employees of an enterprise. The same cut-off point of 50 employees is applied to each wave of the CIS and the PS. A similar scheme is used for the French CIS except that a cut-off point of 500 employees is used in CIS 2 and 3, and a cut-off point of 250 employees is used in CIS 4.

Table 2: Employment, R&D per employee, innovation and sales per employee in each pattern of the unbalanced data for Dutch and French manufacturing: CIS 2, CIS 3, and CIS 4

	France				The Netherlands			
Pattern	110	111	011	Total	110	111	011	Total
# enterprises	532	607	899	2038	580	467	396	1443
% in total	26	30	44	100	40	32	28	100
Employment, sample								
Mean	543	1037	505	720	134	223	275	204
Median	156	660	185	324	68	109	72	85
Employment, population								
Mean	207	709	230	337	84	167	128	124
Median	68	381	68	90	41	85	32	50
Mean interest variables								
Product innovator	0.51	0.72	0.49	0.58	0.56	0.60	0.46	0.55
R&D/employee†	6.78	16.29	9.61	12.50	4.50	5.54	4.84	5.07
Share of innov. sales‡	0.23	0.23	0.20	0.22	0.32	0.30	0.28	0.30
Sales/employee*	214	268	233	242	147	216	160	179

†of continuous R&D performers, in 1000s of euros. ‡of product innovators. *in 1000s of euros.

Table 2 shows, for both countries, the patterns of the unbalanced panel that is used in our analysis. As we explained earlier, R&D is assumed to translate into innovation after one period, which corresponds to four years in this analysis. Hence, an enterprise must be present in at least two consecutive waves of the merged data in order to be included in the analysis. There are 2038 such enterprises in our sample for France and 1443 for The Netherlands. Besides the number of enterprises, Table 2 shows the mean and the median of the number of employees in the sample and in the population,[3] the proportion of product innovators, and the mean value of R&D expenditures per employee, share of innovative sales and total sales per employee. The table shows that French enterprises that are in the balanced panel (pattern 2) are on average significantly (at 5% level) larger in terms of employment and have a significantly larger proportion of product innovators than those of pattern 1 and 3. As for The Netherlands, enterprises that are in the balanced panel have a significantly larger proportion of product innovators and have on average significantly larger sales per employee than those of pattern 1 and 3. When comparing the two countries, the table shows that French enterprises, both in the sample and in the population, are on average significantly

[3]The number of employees in the population is obtained by weighting the number of employees in the sample using the raising factor obtained after correcting for non-response.

larger than the Dutch ones. Furthermore, the former have on average larger sales per employee than the latter. However, the share of innovative sales of Dutch product innovators is on average significantly larger than that of the French counterparts.

3.1 Dependent and explanatory variables

The first dependent variable of the model is binary taking on the value one if an enterprise is a product innovator, and zero otherwise. In the innovation survey, an enterprise is asked whether it has implemented at least one new or improved product during the period under review. A product innovator is an enterprise that has responded positively to such a question. The second dependent variable is the share in total sales accounted for by sales of new or improved products. This variable stems from the CIS and is measured at the end of the period under review. A logit transformation of this variable is used in order to make it lie within the entire set of real numbers.[4] Finally, the third dependent variable is a proxy for labor productivity, namely the ratio of domestic sales over employment. Sales and employment stem from the PS for The Netherlands, and from the CIS for France. A natural log transformation of this proxy is used in the estimation of the model.

The explanatory variables included in the analysis are listed in Table 3. We motivate their inclusion (or not) in the three equations of the model as follows.

Table 3: Specification of the three equations of the panel data CDM model

Regressors	Dependent Variables		
	Product innovator	Share of innov. sales	Sales/employee
R&D/employee[†]	no	yes	no
$D_{\text{non-continuous R\&D}}$	no	yes	no
Employment	yes	yes	yes
Market share	yes	no	no
Technology push	yes	yes	no
Demand pull	yes	yes	no
Share of innov. sales	no	no	yes
Distance to frontier			
D_{Q_2} to D_{Q_4}	yes	yes	no
D_{industry}	yes	yes	yes
D_{time}	yes	yes	yes
Individual effects	yes	yes	yes

[†]for continuous R&D performers.

R&D per capita

The knowledge production function literature identifies R&D as the main factor of knowledge generation (see e.g. Pakes and Griliches, 1980), which in turn is translated into higher innovative

[4]The share of innovative sales takes on the value 1 for innovators that are newly established. It is replaced by 0.99 in the logit transformation.

sales. Ideally, we would have liked to use a stock measure of R&D in the equation of the share of innovative sales. However, the CIS (The Netherlands) and the R&D survey (France) provide us with flow measures of R&D. Hence, we consider in this analysis R&D (per capita, in log) of continuous R&D performers so as to proxy as well as possible a stock measure of R&D. An additional dummy variable for non-continuous R&D performers is included to compensate for the fact that we use positive values of R&D only for continuous R&D performers.

Employment, market share

On the grounds of one of Schumpeter's (1942) hypothesis, we explain the probability of being a product innovator and the share of innovative sales by employment taken to measure firm (absolute) size. It also enters the labor productivity equation as one input to the production function. According to Schumpeter, market share, taken to measure firm relative size and defined as the ratio of the sales of an enterprise over the total sales of the 3-digit industry it belongs to,[5] is also expected to play a significant role in innovation. We do not include market share as a regressor in the equation of the share of innovative sales so as to allow for an exclusion restriction, which is typical of sample selection models (see e.g. Vella, 1998).

Technology push, demand pull

Following Schumpeter (1934) and Schmookler (1966), we explain the incidence and the intensity of product innovation by technology push and demand pull respectively. These two variables are constructed for both countries from the CIS data as follows. Enterprises are asked to quantify the importance of innovation effects on a 0-3 Likert scale. First, two dummy variables of process-oriented effects (i.e. improve flexibility of production, increase its capacity or reduce labor costs, materials or energy per unit output) and product-oriented effects (i.e. increase the range of goods, improve their quality, increase market share or enter new markets) are constructed as taking the value one if the corresponding effects of innovation are deemed very important (value 3 on the scale), and zero otherwise.[6] Then, each enterprise (including the non-innovative ones) is attributed the mean observed value of these dummy variables within the 3-digit industry it belongs to. As a result, our technology push and demand pull variables capture the share of firms within a 3-digit industry for which the above-mentioned process- and product-oriented effects are deemed very important.

[5]Market share actually measures domestic market share, as only domestic sales are considered. Total sales of a 3-digit industry is obtained by adding up the sales of all the firms in our sample that belong to that industry after multiplying them by the appropriate raising factor.

[6]The questions regarding the importance of innovation effects are posed only to firms that are product innovators, process innovators, or firms that have ongoing or abandoned innovation activities. These firms are called innovative, and the dummy variables of process- and product-oriented effects are missing for non-innovative firms.

Share of innovative sales

In addition to being one of the dependent variables of the model, the share of innovative sales enters the labor productivity equation as an additional (knowledge) input to the augmented production function besides the conventional ones (labor and capital) along the lines of Griliches (1979).[7] Since the share of innovative sales is observed to be positive only for product innovators, we use its latent counterpart as a regressor in the labor productivity equation so as to capture the innovation effects of both actual and potential product innovators.

Distance to technological frontier

The notion of technological frontier is mostly used in the macroeconomic literature on growth convergence. Among various testable assumptions is that innovation becomes more important as an economy approaches the world technology frontier (e.g. Acemoglu et al., 2003, 2006). We can mimic this situation at the firm level for a given sector, e.g. manufacturing, in given economies, e.g. France and The Netherlands. In order to achieve that, we identify in each 3-digit industry the enterprise with the largest labor productivity. We then define for each firm a *technology gap* variable as the difference between the largest productivity (within each 3-digit industry) and the productivity of the firm. We proceed by looking at the distribution (within each industry) of the technology gap variable. Finally, we define three dummy variables D_{Q_2}-D_{Q_4} which take on the value one if the technology gap lies respectively between the first (>) and second quartile (≤), the second and the third quartile, and above (>) the third quartile. The dummy variable D_{Q_1}, which takes on the value one if the technology gap lies below or at the first quartile, is used as the reference. Firms for which D_{Q_1} is equal to one are the closest to the technological frontier. If the above-mentioned hypothesis is satisfied, we expect the effects of D_{Q_2}, D_{Q_3} and D_{Q_4} to be negative and statistically significant.

Industry, time and individual effects

In all three equations of the model, we control for industry, time and individual effects. Four categories of industry are distinguished according to the OECD (2007) technology-based classification namely high-tech, medium-high-tech, medium-low-tech, and low-tech. Three dummy variables of industry category, with the high-tech category being the reference, are included in order to capture industry-specific effects such as technological opportunities (it is easier to innovate in certain industries than in others) and intensity of competition (which is expected to be higher in high-tech

[7]Due to a lack of CIS data for both countries and for the three time periods, we could not account for physical capital in the labor productivity equation of this version of the analysis. We plan to do so in the near future by using additional data from other sources.

industries than in low-tech industries) among others. A time dummy variable, with the period 2002-2004 being the reference, is included to capture macroeconomic shocks.[8] Finally, we include individual effects so as to allow for unobserved heterogeneity or (a special form of) serial correlation.

3.2 Descriptive statistics

Table 4 presents descriptive statistics of the dependent and explanatory variables for all firms and for the sub-sample of product innovators. We observe the following patterns. First, the figures are in general significantly larger (on average) for France than for The Netherlands. Three exceptions are the share of innovative sales which is significantly larger for The Netherlands, and market share and R&D per capita that do not differ significantly across the two countries.[9] Secondly and somewhat surprisingly, labor productivity of product innovators in both countries is not statistically different from that of the other firms. One explanation could be that innovation translates into larger productivity only after a certain period of time, while the figures of Table 4 are considered at the same period. Thirdly, product innovators of both countries are significantly larger in terms of employment and have a significantly larger market share than the other firms. The difference in employment of product innovators is particularly pronounced in the case of France. Finally the figures of technology push and demand pull are by construction similar in the sub-sample of product innovators and in the whole sample, and similar across countries.

Table 4: Mean dependent and explanatory variables: Unbalanced panel from Dutch and French CIS 2, CIS 3 and CIS 4

Variable	France		The Netherlands	
	All firms	Product innov.	All firms	Product innov.
Product innovator	0.58	-	0.55	-
Share of innov. sales	0.13	0.22	0.17	0.30
Sales/employee	242.36	243.24	179.00	178.20
R&D/employee†	-	12.50	-	5.07
Employment	720.39	1003.08	204.26	250.12
Market share (%)	1.45	1.82	1.37	1.70
Technology push	0.44	0.45	0.42	0.43
Demand pull	0.76	0.78	0.73	0.74
# observations	4683	2731	3353	1855

†of continuous R&D performers.

Table 5 presents descriptive statistics of the dependent and explanatory variables for each CIS wave of the unbalanced panel. For both countries, the proportion of product innovators and the share of innovative sales of product innovators decrease significantly (on average) between 1994 and

[8]The first period observation is lost because of the lag between R&D and innovation.

[9]The amount of R&D per capita of French continuous R&D performers is larger than that of Dutch continuous R&D performers. However, that variable has a very large (not reported) standard error in the case of France, which makes the difference across the two countries statistically insignificant.

2004, while sales per capita increase significantly over that period. The increase in the two measures of innovation for France between the last two periods is not large enough, albeit significant for the share of innovative sales, to offset the large decrease that occurs between the first two periods. Furthermore, R&D per capita of continuous R&D performers increases significantly (on average) for The Netherlands,[10] and market share increases significantly for both countries between 1994 and 2004, where the increase in the latter measure mainly occurs between the last (first) two periods for France (The Netherlands). The two measures of technology push and demand pull are similar across countries and over time. Finally, there is a decrease in the mean employment, albeit insignificant at 5% level, for France and a significant increase for The Netherlands between 1994 and 2004.

Table 5: Mean dependent and explanatory variables for each CIS of the unbalanced panel, France and The Netherlands

Variable	France			The Netherlands		
	1994-1996	1998-2000	2002-2004	1994-1996	1998-2000	2002-2004
Product innovator	0.65	0.55	0.57	0.63	0.57	0.43
Share of innov. sales†	0.31	0.15	0.23	0.33	0.32	0.22
Sales/employee	170.11	257.39	276.68	144.91	166.86	240.65
R&D/employee‡	8.49	8.16	21.19	3.95	5.49	5.92
Employment	796.37	711.42	675.05	171.49	210.08	234.30
Market share (%)	1.29	1.38	1.68	1.11	1.42	1.60
Technology push	0.44	0.44	0.44	0.43	0.42	0.42
Demand pull	0.76	0.76	0.76	0.72	0.73	0.73
# enterprises	1139	2038	1506	1047	1443	863

†of product innovators. ‡of continuous R&D performers.

4 Model

Let y^*_{jit} $(j = 1, 2, 3)$ denote three latent variables representing respectively firm's i incentive to innovate, its share of innovative sales (innovation output) and its productivity at period t $(t = 1, ..., T_i)$. These latent variables can be written as functions of (vectors of) strictly exogenous explanatory variables $\mathbf{x}_{jit}$, time-invariant individual effects α_{ji} possibly correlated with the exogenous regressors and other time-varying unobserved variables summarized in idiosyncratic errors ϵ_{jit}. Productivity

[10]The mean of R&D/employee for France is much larger in the last period than in the first two periods. However, its standard error (not reported) is so large that the mean difference, say, between the first and the last period, is not statistically significant.

is also assumed to be explained by the share of innovative sales. Formally we write

$$y_{1it}^* = \boldsymbol{\beta}_1' \mathbf{x}_{1it} + \alpha_{1i} + \epsilon_{1it}, \tag{4.0.1a}$$

$$y_{2it}^* = \boldsymbol{\beta}_2' \mathbf{x}_{2it} + \alpha_{2i} + \epsilon_{2it}, \tag{4.0.1b}$$

$$y_{3it}^* = \boldsymbol{\beta}_3' \mathbf{x}_{3it} + \gamma y_{2it}^* + \alpha_{3i} + \epsilon_{3it}, \tag{4.0.1c}$$

where γ captures the relationship between productivity and innovation output and is to be estimated together with $\boldsymbol{\beta}_j$. The individual effects and idiosyncratic errors are allowed to be correlated across equations, but mutually independent within each equation.

Taken separately, equation (4.0.1b) can be seen as a knowledge production function where the main input to the share of innovative sales is R&D, and equation (4.0.1c) models an augmented production function with physical capital, labor and knowledge capital as the main inputs. Taken jointly, equations (4.0.1a)-(4.0.1c) is a type of simultaneous-equation model where one of the regressors, the share of innovative sales, is endogenous and simultaneously determined with the other explained variables. The observed counterparts to equations (4.0.1a)-(4.0.1c) are written as

$$(y_{1it}, y_{2it}, y_{3it}) = \begin{cases} (0, 0, y_{3it}^*) & \text{if } y_{1it}^* \leq 0 \\ (1, y_{2it}^*, y_{3it}^*) & \text{if } y_{1it}^* > 0 \end{cases} . \tag{4.0.2}$$

If the incentive to innovate of firm i crosses a certain threshold (say 0) at period t, the firm is observed to be a product innovator and its share of innovative sales is observed to be positive. If on the contrary the innovation incentive does not cross the threshold, the firm is observed to be a non-product innovator and its share of innovative sales is zero. In the two cases, we observe the (labor) productivity of the firm.

Two variants of this model can be considered depending on the assumptions we make about the individual effects. In the first variant, we assume that the time-invariant unobserved heterogeneity of the two equations of incidence and intensity of innovation (eqs. (4.0.1a) and (4.0.1b)) are proxied by the same observed and unobserved variables that have a different effect on the dependent variables. We will be referring to this assumption as *common factor* individual effects. As we will see, this variant of the model has the virtue of being easier to estimate than the (later-described) second variant, which induces the cost of assuming a correlation of one between α_{1i} and α_{2i}. The second variant relaxes the assumption of common factor individual effects and estimates the correlation between the individual effects. We now describe both variants and their full information maximum likelihood (FIML) estimation.

4.1 Variant 1

The model with common factor individual effects assumes

$$\alpha_{1i} = \delta_1 \eta_i, \;\; \alpha_{2i} = \delta_2 \eta_i, \tag{4.1.1}$$

where η_i represents the time-invariant unobserved heterogeneity in the two equations with different effects δ_1 and δ_2, is allowed to be correlated with α_{3i} but is independent of ϵ_{jit}. For identification purpose, we use the normalization $\delta_1 = 1$.

In order to estimate variant 1 we replace α_{1i} and α_{2i} by their expression into equations (4.0.1a) and (4.0.1b) and make the following assumptions: $(\eta_i, \alpha_{3i})'$ is identically and independently distributed (*iid*) across individuals and over time following a normal distribution with mean zero and covariance matrix

$$\Sigma_{\eta\alpha_3} = \begin{pmatrix} \sigma^2_\eta & \\ \rho_{\eta\alpha_3}\sigma_\eta\sigma_{\alpha_3} & \sigma^2_{\alpha_3} \end{pmatrix} \tag{4.1.2}$$

and, conditionnally on η_i and α_{3i}, $(\epsilon_{1it}, \epsilon_{2it}, \epsilon_{3it})'$ is *iid* normally distributed with mean zero and covariance matrix

$$\Sigma_{\boldsymbol{\epsilon}} = \begin{pmatrix} 1 & & \\ \rho_{12}\sigma_{\epsilon_2} & \sigma^2_{\epsilon_2} & \\ \rho_{13}\sigma_{\epsilon_3} & \rho_{23}\sigma_{\epsilon_2}\sigma_{\epsilon_3} & \sigma^2_{\epsilon_3} \end{pmatrix}. \tag{4.1.3}$$

The parameters of equations (4.0.1a)-(4.0.1c) and (4.1.1), and those of equations (4.1.2) and (4.1.3) are the *structural* parameters of variant 1 of the model and are of interest. The *reduced-form* latent equations of the model are given by equations (4.0.1a) and (4.0.1b), and

$$y^*_{3it} = \boldsymbol{\beta}'_3\mathbf{x}_{3it} + \underbrace{\gamma\boldsymbol{\beta}'_2}_{\underline{\boldsymbol{\beta}}'_2}\mathbf{x}_{2it} + \underbrace{\gamma\delta_2\eta_i + \alpha_{3i}}_{\underline{\alpha}_{3i}} + \underbrace{\gamma\epsilon_{2it} + \epsilon_{3it}}_{\underline{\epsilon}_{3it}} \tag{4.1.4}$$

after substituting into equation (4.0.1c) the right-hand side of equation (4.0.1b) for y^*_{2it}. If $\mathbf{x}_{2it}$ and $\mathbf{x}_{3it}$ have common components, their associated coefficients are not identified separately from each other in equation (4.1.4). Because of the normal distribution assumptions, $\underline{\alpha}_{3i}$ and $\underline{\epsilon}_{3it}$ are also normally distributed with mean zero and variance $\sigma^2_{\underline{\alpha}_3}$ and $\sigma^2_{\underline{\epsilon}_3}$ respectively. Hence, the covariance matrix of the reduced-form individual effects and idiosyncratic errors can be written as

$$\Sigma_{\eta\underline{\alpha}_3} = \begin{pmatrix} \sigma^2_\eta & \\ \rho_{\eta\underline{\alpha}_3}\sigma_\eta\sigma_{\underline{\alpha}_3} & \sigma^2_{\underline{\alpha}_3} \end{pmatrix} ; \; \Sigma_{\underline{\boldsymbol{\epsilon}}} = \begin{pmatrix} 1 & & \\ \rho_{12}\sigma_{\epsilon_2} & \sigma^2_{\epsilon_2} & \\ \underline{\rho}_{13}\sigma_{\underline{\epsilon}_3} & \underline{\rho}_{23}\sigma_{\epsilon_2}\sigma_{\underline{\epsilon}_3} & \sigma^2_{\underline{\epsilon}_3} \end{pmatrix} \tag{4.1.5}$$

where $\rho_{\eta\underline{\alpha}_3}$ denotes the correlation between the common factor individual effects (η_i) and the reduced-form individual effects ($\underline{\alpha}_{3i}$), and $\underline{\rho}_{k3}$ ($k = 1, 2$) denotes the correlations between the idiosyncratic errors ϵ_{kit} and the reduced-form idiosyncratic error $\underline{\epsilon}_{3it}$. The variance of the reduced-form individual effects as well as their correlation with the common factor individual effects are given by

$$\sigma^2_{\underline{\alpha}_3} = \gamma^2\delta_2^2\sigma^2_\eta + \sigma^2_{\alpha_3} + 2\gamma\delta_2\rho_{\eta\alpha_3}\sigma_\eta\sigma_{\alpha_3},\ \rho_{\eta\underline{\alpha}_3} = \frac{\gamma\delta_2\sigma_\eta + \rho_{\eta\alpha_3}\sigma_{\alpha_3}}{\sigma_{\underline{\alpha}_3}}. \tag{4.1.6}$$

Similarly, the variance of the reduced-form idiosyncratic error and its correlations with ϵ_{kit} can be written as

$$\sigma^2_{\underline{\epsilon}_3} = \gamma^2\sigma^2_{\epsilon_2} + \sigma^2_{\epsilon_3} + 2\gamma\rho_{23}\sigma_{\epsilon_2}\sigma_{\epsilon_3},\ \underline{\rho}_{13} = \frac{\gamma\rho_{12}\sigma_{\epsilon_2} + \rho_{13}\sigma_{\epsilon_3}}{\sigma_{\underline{\epsilon}_3}},\ \underline{\rho}_{23} = \frac{\gamma\sigma_{\epsilon_2} + \rho_{23}\sigma_{\epsilon_3}}{\sigma_{\underline{\epsilon}_3}}. \tag{4.1.7}$$

We can estimate the structural parameters of variant 1 of the model by FIML which consists in estimating the reduced-form model by maximum likelihood using the additional information of equations (4.1.6) and (4.1.7). The likelihood function of the reduced form of variant 1 of the model can be written as

$$L_{1i} = \int_{\eta_i}\int_{\underline{\alpha}_{3i}} L_{1i}(...|..., \eta_i, \underline{\alpha}_{3i}) g_2(\eta_i, \underline{\alpha}_{3i}|...) d\eta_i d\underline{\alpha}_{3i}, \tag{4.1.8}$$

where $g_2(\eta_i, \underline{\alpha}_{3i}|...)$ denotes the (bivariate) distribution of the reduced-form individual effects given the regressors. We can show that $L_{1i}(...|..., \eta_i, \underline{\alpha}_{3i})$ is equal to

$$\begin{aligned}
&\prod_{i=1}^{T_i} \frac{1}{\sigma_{\underline{\epsilon}_3}}\phi_1\left(\frac{y_{3it} - N_{3it} - \underline{\alpha}_{3i}}{\sigma_{\underline{\epsilon}_3}}\right)\left[\Phi_1\left(\frac{-N_{1it} - \eta_i - \frac{\underline{\rho}_{13}}{\sigma_{\underline{\epsilon}_3}}(y_{3it} - N_{3it} - \underline{\alpha}_{3i})}{\sqrt{1 - \underline{\rho}^2_{13}}}\right)\right]^{1-y_{1it}} \\
&\left[\frac{1}{\sigma_{\epsilon_2}\sqrt{1 - \underline{\rho}^2_{23}}}\phi_1\left(\frac{y_{2it} - N_{2it} - \delta_2\eta_i - \underline{\rho}_{23}\frac{\sigma_{\epsilon_2}}{\sigma_{\underline{\epsilon}_3}}(y_{3it} - N_{3it} - \underline{\alpha}_{3i})}{\sigma_{\epsilon_2}\sqrt{1 - \underline{\rho}^2_{23}}}\right)\right. \\
&\left.\Phi_1\left(\frac{N_{1it} + \eta_i + \frac{\underline{\rho}_{12.3}}{\sigma_{\epsilon_2}}(y_{2it} - N_{2it} - \delta_2\eta_i) + \frac{\underline{\rho}_{13.2}}{\sigma_{\underline{\epsilon}_3}}(y_{3it} - N_{3it} - \underline{\alpha}_{3i})}{\sqrt{1 - \underline{R}^2_{1.23}}}\right)\right]^{y_{1it}},
\end{aligned} \tag{4.1.9}$$

where ϕ_1 and Φ_1 denote respectively the univariate standard normal pdf and cdf, and

$$N_{1it} \equiv \boldsymbol{\beta}_1'\mathbf{x}_{1it},\ N_{2it} \equiv \boldsymbol{\beta}_2'\mathbf{x}_{2it},\ N_{3it} \equiv \boldsymbol{\beta}_3'\mathbf{x}_{3it} + \underline{\boldsymbol{\beta}}_2'\mathbf{x}_{2it}, \tag{4.1.10}$$

and

$$\underline{R}^2_{1.23} \equiv \frac{\rho^2_{12} + \underline{\rho}^2_{13} - 2\rho_{12}\underline{\rho}_{13}\underline{\rho}_{23}}{1 - \underline{\rho}^2_{23}},\ \underline{\rho}_{12.3} \equiv \frac{\rho_{12} - \underline{\rho}_{13}\underline{\rho}_{23}}{1 - \underline{\rho}^2_{23}},\ \underline{\rho}_{13.2} \equiv \frac{\underline{\rho}_{13} - \rho_{12}\underline{\rho}_{23}}{1 - \underline{\rho}^2_{23}}.$$

To "integrate out" the reduced-form individual effects, we use their bivariate normal density and *two-step* Gauss-Hermite (G-H) quadrature as follows. Equation (4.1.8) can be written as

$$L_{1i} = (2\pi\sigma_\eta\sigma_{\underline{\alpha}_3})^{-1}\Big(1-\rho^2_{\eta\underline{\alpha}_3}\Big)^{\frac{-1}{2}}\int_{\underline{\alpha}_{3i}} e^{\frac{-\underline{\alpha}^2_{3i}}{2\sigma^2_{\underline{\alpha}_3}(1-\rho^2_{\eta\underline{\alpha}_3})}}\prod_{i=1}^{T_i}\frac{1}{\sigma_{\underline{\epsilon}_3}}\phi_1\Big(\frac{y_{3it}-N_{3it}-\underline{\alpha}_{3i}}{\sigma_{\underline{\epsilon}_3}}\Big)F(\underline{\alpha}_{3i}|...)d\underline{\alpha}_{3i}, \tag{4.1.11}$$

where

$$F(\underline{\alpha}_{3i}|...) \equiv \int_{\eta_i} e^{\frac{-\eta_i^2}{2\sigma^2_\eta(1-\rho^2_{\eta\underline{\alpha}_3})}} e^{\frac{\rho^2_{\eta\underline{\alpha}_3}\frac{\eta_i}{\sigma_\eta}\frac{\underline{\alpha}_{3i}}{\sigma_{\underline{\alpha}_3}}}{(1-\rho^2_{\eta\underline{\alpha}_3})}}\prod_{i=1}^{T_i}\Bigg[\Phi_1\Bigg(\frac{-N_{1it}-\eta_i-\frac{\underline{\rho}_{13}}{\sigma_{\underline{\epsilon}_3}}(y_{3it}-N_{3it}-\underline{\alpha}_{3i})}{\sqrt{1-\underline{\rho}^2_{13}}}\Bigg)\Bigg]^{1-y_{1it}}$$
$$\Bigg[\frac{1}{\sigma_{\epsilon_2}\sqrt{1-\underline{\rho}^2_{23}}}\phi_1\Bigg(\frac{y_{2it}-N_{2it}-\delta_2\eta_i-\underline{\rho}_{23}\frac{\sigma_{\epsilon_2}}{\sigma_{\underline{\epsilon}_3}}(y_{3it}-N_{3it}-\underline{\alpha}_{3i})}{\sigma_{\epsilon_2}\sqrt{1-\underline{\rho}^2_{23}}}\Bigg) \tag{4.1.12}$$
$$\Phi_1\Bigg(\frac{N_{1it}+\eta_i+\frac{\underline{\rho}_{12.3}}{\sigma_{\epsilon_2}}(y_{2it}-N_{2it}-\delta_2\eta_i)+\frac{\underline{\rho}_{13.2}}{\sigma_{\underline{\epsilon}_3}}(y_{3it}-N_{3it}-\underline{\alpha}_{3i})}{\sqrt{1-\underline{R}^2_{1.23}}}\Bigg)\Bigg]^{y_{1it}} d\eta_i.$$

In the first step, we evaluate $F(\underline{\alpha}_{3i}|...)$ using the G-H approximation by making the variable change $z_{1i} = \frac{\eta_i}{\sigma_\eta\sqrt{2(1-\rho^2_{\eta\underline{\alpha}_3})}}$. We then replace the evaluated expression of $F(\underline{\alpha}_{3i}|...)$ into equation (4.1.11) and use again the G-H quadrature by making the variable change $\underline{z}_{2i} = \frac{\underline{\alpha}_{3i}}{\sigma_{\underline{\alpha}_3}\sqrt{2(1-\rho^2_{\eta\underline{\alpha}_3})}}$. The final expression of the likelihood of the reduced form of the model is written as

$$L_{1i} \simeq \frac{\sqrt{1-\rho^2_{\eta\underline{\alpha}_3}}}{\pi}\sum_{p=1}^{P} w_p\prod_{i=1}^{T_i}\frac{1}{\sigma_{\underline{\epsilon}_3}}\phi_1\left(\frac{y_{3it}-N_{3it}-a_p(...)}{\sigma_{\underline{\epsilon}_3}}\right)$$
$$\sum_{m=1}^{M} w_m e^{2\rho_{\eta\underline{\alpha}_3}a_m a_p}\prod_{i=1}^{T_i}\left[\Phi_1\left(\frac{-N_{1it}-a_m(...)-\frac{\underline{\rho}_{13}}{\sigma_{\underline{\epsilon}_3}}(y_{3it}-N_{3it}-a_p(...))}{\sqrt{1-\underline{\rho}^2_{13}}}\right)\right]^{1-y_{1it}}$$
$$\left[\frac{1}{\sigma_{\epsilon_2}\sqrt{1-\underline{\rho}^2_{23}}}\phi_1\left(\frac{y_{2it}-N_{2it}-\delta_2 a_m(...)-\underline{\rho}_{23}\frac{\sigma_{\epsilon_2}}{\sigma_{\underline{\epsilon}_3}}(y_{3it}-N_{3it}-a_p(...))}{\sigma_{\epsilon_2}\sqrt{1-\underline{\rho}^2_{23}}}\right)\right. \tag{4.1.13}$$
$$\left.\Phi_1\left(\frac{N_{1it}+a_m(...)+\frac{\underline{\rho}_{12.3}}{\sigma_{\epsilon_2}}(y_{2it}-N_{2it}-\delta_2 a_m(...))+\frac{\underline{\rho}_{13.2}}{\sigma_{\underline{\epsilon}_3}}(y_{3it}-N_{3it}-a_p(...))}{\sqrt{1-\underline{R}^2_{1.23}}}\right)\right]^{y_{1it}},$$

where w_m, w_p, a_m and a_p are respectively the weights and abscissas of the first- and second-stage G-H integration with M and P being the first- and second-stage total number of integration points, and $a_m(...) = a_m\sigma_\eta\sqrt{2(1-\rho^2_{\eta\underline{\alpha}_3})}$ and $a_p(...) = a_p\sigma_{\underline{\alpha}_3}\sqrt{2(1-\rho^2_{\eta\underline{\alpha}_3})}$. Tables with numerical values of the weights and abscissas of the G-H quadrature are formulated in mathematical textbooks (e.g. Abramowitz and Stegun, 1964). We obtain the log-likelihood as

$$L_1 = \sum_{i=1}^{N}\ln L_{1i} \tag{4.1.14}$$

which is maximized using standard numerical techniques (e.g. Newton-Raphson) to yield ML estimates of the reduced form of variant 1 of the panel data CDM model. The FIML estimates of the structural parameters of the model are obtained by maximizing L_1 and replacing (into equation (4.1.13)) $\underline{\boldsymbol{\beta}}'_2$ by $\gamma\boldsymbol{\beta}'_2$, $\sigma_{\underline{\alpha}_3}$ and $\rho_{\eta\underline{\alpha}_3}$ by their expressions given in equation (4.1.6), and $\sigma_{\underline{\epsilon}_3}$, $\underline{\rho}_{13}$ and $\underline{\rho}_{23}$ by their expressions given in equation (4.1.7).

4.2 Variant 2

This variant of the model relaxes the common factor individual effects assumption. More specifically, we consider three different components of individual effects α_{1i}, α_{2i} and α_{3i} that are allowed to be correlated across equations. Again, conditionally on the three components of individual effects, $(\epsilon_{1it}, \epsilon_{2it}, \epsilon_{3it})'$ is *iid* normally distributed with mean zero and covariance matrix given in equation (4.1.3). The individual effects are now assumed to have a trivariate normal distribution with mean zero and covariance matrix

$$\Sigma_{\boldsymbol{\alpha}} = \begin{pmatrix} \sigma^2_{\alpha_1} & & \\ \rho_{\alpha_1\alpha_2}\sigma_{\alpha_1}\sigma_{\alpha_2} & \sigma^2_{\alpha_2} & \\ \rho_{\alpha_1\alpha_3}\sigma_{\alpha_1}\sigma_{\alpha_3} & \rho_{\alpha_2\alpha_3}\sigma_{\alpha_2}\sigma_{\alpha_3} & \sigma^2_{\alpha_3} \end{pmatrix}. \tag{4.2.1}$$

The structural parameters of variant 2 of the model are now listed in equations (4.0.1a)-(4.0.1c), (4.1.3) and (4.2.1). The counterpart to equation (4.1.4) is given by

$$y^*_{3it} = \boldsymbol{\beta}'_3\mathbf{x}_{3it} + \underbrace{\gamma\boldsymbol{\beta}'_2}_{\underline{\boldsymbol{\beta}}'_2}\mathbf{x}_{2it} + \underbrace{\gamma\alpha_{2i} + \alpha_{3i}}_{\underline{\underline{\alpha}}_{3i}} + \underbrace{\gamma\epsilon_{2it} + \epsilon_{3it}}_{\underline{\epsilon}_{3it}}. \tag{4.2.2}$$

The covariance matrix of the reduced-form idiosyncratic errors remains the same, hence the expressions of $\sigma^2_{\underline{\epsilon}_3}$, $\underline{\rho}_{13}$ and $\underline{\rho}_{23}$ are given in equation (4.1.7). As for the reduced-form individual effects, the covariance matrix now becomes

$$\Sigma_{\underline{\underline{\boldsymbol{\alpha}}}} = \begin{pmatrix} \sigma^2_{\alpha_1} & & \\ \rho_{\alpha_1\alpha_2}\sigma_{\alpha_1}\sigma_{\alpha_2} & \sigma^2_{\alpha_2} & \\ \rho_{\alpha_1\underline{\underline{\alpha}}_3}\sigma_{\alpha_1}\sigma_{\underline{\underline{\alpha}}_3} & \rho_{\alpha_2\underline{\underline{\alpha}}_3}\sigma_{\alpha_2}\sigma_{\underline{\underline{\alpha}}_3} & \sigma^2_{\underline{\underline{\alpha}}_3} \end{pmatrix},$$

where $\sigma^2_{\underline{\underline{\alpha}}_3}$, $\rho_{\alpha_1\underline{\underline{\alpha}}_3}$ and $\rho_{\alpha_2\underline{\underline{\alpha}}_3}$ are given by

$$\sigma^2_{\underline{\underline{\alpha}}_3} = \gamma^2\sigma^2_{\alpha_2} + \sigma^2_{\alpha_3} + 2\gamma\rho_{\alpha_2\alpha_3}\sigma_{\alpha_2}\sigma_{\alpha_3}, \;\; \rho_{\alpha_1\underline{\underline{\alpha}}_3} = \frac{\gamma\rho_{\alpha_1\alpha_2}\sigma_{\alpha_2} + \rho_{\alpha_1\alpha_3}\sigma_{\alpha_3}}{\sigma_{\underline{\underline{\alpha}}_3}}, \;\; \rho_{\alpha_2\underline{\underline{\alpha}}_3} = \frac{\gamma\sigma_{\alpha_2} + \rho_{\alpha_2\alpha_3}\sigma_{\alpha_3}}{\sigma_{\underline{\underline{\alpha}}_3}}. \tag{4.2.3}$$

The likelihood function of the reduced form of variant 2 of the model is written as

$$L_{2i} = \int_{\alpha_{1i}} \int_{\alpha_{2i}} \int_{\underline{\underline{\alpha}}_{3i}} L_{2i}(...|..., \alpha_{1i}, \alpha_{2i}, \underline{\underline{\alpha}}_{3i}) g_3(\alpha_{1i}, \alpha_{2i}, \underline{\underline{\alpha}}_{3i}|...) d\alpha_{1i} d\alpha_{2i} d\underline{\underline{\alpha}}_{3i}, \tag{4.2.4}$$

where $g_3(\alpha_{1i}, \alpha_{2i}, \underline{\underline{\alpha}}_{3i}|...)$ denotes the (trivariate) distribution of the reduced-form individual effects given the regressors. The conditional individual likelihood $L_{2i}(...|..., \alpha_{1i}, \alpha_{2i}, \underline{\underline{\alpha}}_{3i})$, which is the counterpart to equation (4.1.9) is given by

$$\begin{aligned}
&\prod_{i=1}^{T_i} \frac{1}{\sigma_{\underline{\epsilon}_3}} \phi_1 \left(\frac{y_{3it} - N_{3it} - \underline{\underline{\alpha}}_{3i}}{\sigma_{\underline{\epsilon}_3}} \right) \left[\Phi_1 \left(\frac{-N_{1it} - \alpha_{1i} - \frac{\underline{\rho}_{13}}{\sigma_{\underline{\epsilon}_3}} (y_{3it} - N_{3it} - \underline{\underline{\alpha}}_{3i})}{\sqrt{1 - \underline{\rho}_{13}^2}} \right) \right]^{1-y_{1it}} \\
&\left[\frac{1}{\sigma_{\epsilon_2} \sqrt{1 - \underline{\rho}_{23}^2}} \phi_1 \left(\frac{y_{2it} - N_{2it} - \alpha_{2i} - \underline{\rho}_{23} \frac{\sigma_{\epsilon_2}}{\sigma_{\underline{\epsilon}_3}} (y_{3it} - N_{3it} - \underline{\underline{\alpha}}_{3i})}{\sigma_{\epsilon_2} \sqrt{1 - \underline{\rho}_{23}^2}} \right) \right. \\
&\left. \Phi_1 \left(\frac{N_{1it} + \alpha_{1i} + \frac{\underline{\rho}_{12.3}}{\sigma_{\epsilon_2}} (y_{2it} - N_{2it} - \alpha_{2i}) + \frac{\underline{\rho}_{13.2}}{\sigma_{\underline{\epsilon}_3}} (y_{3it} - N_{3it} - \underline{\underline{\alpha}}_{3i})}{\sqrt{1 - \underline{R}_{1.23}^2}} \right) \right]^{y_{1it}},
\end{aligned} \tag{4.2.5}$$

with N_{jit} $(j = 1, 2, 3)$ defined in equation (4.1.10).

In order to obtain the unconditional likelihood, we need to evaluate the triple integral of equation (4.2.4), which is achieved by using *three-step* G-H. The principle is similar to that described in equations (4.1.11)-(4.1.13) and works as follows. Using the expression of the trivariate normal density (see Appendix B) and equation (4.2.5) we can rewrite equation (4.2.4) as

$$L_{2i} = \Lambda \int_{\underline{\underline{\alpha}}_{3i}} e^{-A_{33} \frac{\underline{\underline{\alpha}}_{3i}^2}{2\sigma_{\underline{\underline{\alpha}}_3}^2}} \prod_{i=1}^{T_i} \frac{1}{\sigma_{\underline{\epsilon}_3}} \phi_1 \left(\frac{y_{3it} - N_{3it} - \underline{\underline{\alpha}}_{3i}}{\sigma_{\underline{\epsilon}_3}} \right) G(\underline{\underline{\alpha}}_{3i}|...) d\underline{\underline{\alpha}}_{3i}, \tag{4.2.6}$$

where the function $G(\underline{\underline{\alpha}}_{3i}|...)$ is defined as

$$G(\underline{\underline{\alpha}}_{3i}|...) \equiv \int_{\alpha_{2i}} e^{-A_{22} \frac{\alpha_{2i}^2}{2\sigma_{\alpha_2}^2}} e^{-A_{23} \frac{\alpha_{2i}}{\sigma_{\alpha_2}} \frac{\underline{\underline{\alpha}}_{3i}}{\sigma_{\underline{\underline{\alpha}}_3}}} \prod_{i=1}^{T_i} \left[\frac{\phi_1 \left(\frac{y_{2it} - N_{2it} - \alpha_{2i} - \underline{\rho}_{23} \frac{\sigma_{\epsilon_2}}{\sigma_{\underline{\epsilon}_3}} (y_{3it} - N_{3it} - \underline{\underline{\alpha}}_{3i})}{\sigma_{\epsilon_2} \sqrt{1 - \underline{\rho}_{23}^2}} \right)}{\sigma_{\epsilon_2} \sqrt{1 - \underline{\rho}_{23}^2}} \right]^{y_{1it}} H(\alpha_{2i}, \underline{\underline{\alpha}}_{3i}|...) d\alpha_{2i}, \tag{4.2.7}$$

and $H(\alpha_{2i}, \underline{\underline{\alpha}}_{3i}|...)$ is defined as

$$\begin{aligned}
H(\alpha_{2i}, \underline{\underline{\alpha}}_{3i}|...) \equiv & \int_{\alpha_{1i}} e^{-A_{11} \frac{\alpha_{1i}^2}{2\sigma_{\alpha_1}^2}} e^{\frac{-\alpha_{1i}}{\sigma_{\alpha_1}} \left(A_{12} \frac{\alpha_{2i}}{\sigma_{\alpha_2}} + A_{13} \frac{\underline{\underline{\alpha}}_{3i}}{\sigma_{\underline{\underline{\alpha}}_3}} \right)} \prod_{i=1}^{T_i} \left[\Phi_1 \left(\frac{-N_{1it} - \alpha_{1i} - \frac{\underline{\rho}_{13}}{\sigma_{\underline{\epsilon}_3}} (y_{3it} - N_{3it} - \underline{\underline{\alpha}}_{3i})}{\sqrt{1 - \underline{\rho}_{13}^2}} \right) \right]^{1-y_{1it}} \\
& \left[\Phi_1 \left(\frac{N_{1it} + \alpha_{1i} + \frac{\underline{\rho}_{12.3}}{\sigma_{\epsilon_2}} (y_{2it} - N_{2it} - \alpha_{2i}) + \frac{\underline{\rho}_{13.2}}{\sigma_{\underline{\epsilon}_3}} (y_{3it} - N_{3it} - \underline{\underline{\alpha}}_{3i})}{\sqrt{1 - \underline{R}_{1.23}^2}} \right) \right]^{y_{1it}} d\alpha_{1i}.
\end{aligned} \tag{4.2.8}$$

The expressions of Λ and A_{jk} $(j, k = 1, 2, 3;\ A_{jk} = A_{kj})$ are given respectively in equations (B.0.12) and (B.0.13) of Appendix B. The first step of the approach consists in evaluating the integral of equation (4.2.8) using the G-H quadrature by making the variable change $z_{1i} = \frac{\alpha_{1i}\sqrt{A_{11}}}{\sigma_{\alpha_1}\sqrt{2}}$. We then replace the evaluated expression of $H(\alpha_{2i}, \underline{\underline{\alpha}}_{3i}|...)$ into equation (4.2.7) and apply again the G-H quadrature by making the variable change $z_{2i} = \frac{\alpha_{2i}\sqrt{A_{22}}}{\sigma_{\alpha_2}\sqrt{2}}$. The final step consists in replacing the evaluated expression of $G(\underline{\underline{\alpha}}_{3i}|...)$ into equation (4.2.6) and apply again the G-H quadrature by making the variable change $\underline{\underline{z}}_{3i} = \frac{\underline{\underline{\alpha}}_{3i}\sqrt{A_{33}}}{\sigma_{\underline{\underline{\alpha}}_3}\sqrt{2}}$. The resulting evaluated individual unconditional likelihood of the reduced form of variant 2 is given by

$$
\begin{aligned}
L_{2i} \simeq & \Delta\pi^{\frac{-3}{2}} \left[(1-\rho^2_{\alpha_1\alpha_2})(1-\rho^2_{\alpha_1\underline{\underline{\alpha}}_3})(1-\rho^2_{\alpha_2\underline{\underline{\alpha}}_3})\right]^{\frac{-1}{2}} \sum_{q=1}^{Q} w_q \prod_{i=1}^{T_i} \frac{1}{\sigma_{\underline{\epsilon}_3}} \phi_1\left(\frac{y_{3it}-N_{3it}-a_q(...)}{\sigma_{\underline{\epsilon}_3}}\right) \\
& \sum_{p=1}^{P} w_p e^{\frac{-2A_{23}a_p a_q}{\sqrt{A_{22}A_{33}}}} \prod_{i=1}^{T_i} \left[\frac{\phi_1\left(\frac{y_{2it}-N_{2it}-a_p(...)-\underline{\rho}_{23}\frac{\sigma_{\epsilon_2}}{\sigma_{\underline{\epsilon}_3}}(y_{3it}-N_{3it}-a_q(...))}{\sigma_{\epsilon_2}\sqrt{1-\underline{\rho}^2_{23}}}\right)}{\sigma_{\epsilon_2}\sqrt{1-\underline{\rho}^2_{23}}}\right]^{y_{1it}} \\
& \sum_{m=1}^{P} w_m e^{\frac{-2a_m}{\sqrt{A_{11}}}\left(\frac{a_p A_{12}}{\sqrt{A_{22}}}+\frac{a_q A_{13}}{\sqrt{A_{33}}}\right)} \prod_{i=1}^{T_i} \left[\Phi_1\left(\frac{-N_{1it}-a_m(...)-\frac{\rho_{13}}{\sigma_{\underline{\epsilon}_3}}(y_{3it}-N_{3it}-a_q(...))}{\sqrt{1-\underline{\rho}^2_{13}}}\right)\right]^{1-y_{1it}} \\
& \left[\Phi_1\left(\frac{N_{1it}+a_m(...)+\frac{\rho_{12.3}}{\sigma_{\epsilon_2}}(y_{2it}-N_{2it}-a_p(...))+\frac{\rho_{13.2}}{\sigma_{\underline{\epsilon}_3}}(y_{3it}-N_{3it}-a_q(...))}{\sqrt{1-\underline{R}^2_{1.23}}}\right)\right]^{y_{1it}},
\end{aligned}
\tag{4.2.9}
$$

where w_m, w_p, w_q, a_m, a_p and a_q are respectively the weights and abscissas of the first-, second- and third-stage G-H integration with M, P and Q being the first-, second- and the third-stage total number of integration points, and $a_m(...) = \frac{a_m\sigma_{\alpha_1}\sqrt{2}}{\sqrt{A_{11}}}$, $a_p(...) = \frac{a_p\sigma_{\alpha_2}\sqrt{2}}{\sqrt{A_{22}}}$ and $a_q(...) = \frac{a_q\sigma_{\underline{\underline{\alpha}}_3}\sqrt{2}}{\sqrt{A_{33}}}$. Like variant 1, we obtain the log-likelihood as

$$
L_2 = \sum_{i=1}^{N} \ln L_{2i} \tag{4.2.10}
$$

which is maximized using standard numerical techniques to yield ML estimates of the reduced form of variant 2. The FIML estimates of the structural parameters of variant 2 are obtained by maximizing L_2 and replacing (into equation (4.2.9) $\underline{\boldsymbol{\beta}}'_2$ by $\gamma\boldsymbol{\beta}'_2$, the reduced-form components of the idiosyncratic errors by their expressions given in equation (4.1.7) and the reduced-form components of the individual effects by their expressions given in equation (4.2.3).

5 Results

Table 6 shows the estimation results of variant 1 of the model when no restriction is imposed on the cross-equation correlations and Table 7 shows the results when the restriction $\rho_{13} = \rho_{23} = \rho_{\eta\alpha_3} = 0$

is imposed.

Table 6: FIML estimates of variant 1 of the panel data CDM model: Unbalanced panel from Dutch and French CIS 2, CIS 3 and CIS 4‡

Variable	Coefficient	(Std. Err.)	Coefficient	(Std. Err.)
	France		The Netherlands	
	Innovation incidence, product innovator			
Employment (log)	0.349**	(0.040)	0.325**	(0.050)
Market share (log)	0.021	(0.027)	0.046	(0.032)
Technology push	0.547†	(0.292)	0.096	(0.498)
Demand pull	0.515†	(0.284)	0.544	(0.453)
Distance to frontier				
D_{Q_2}	-0.051	(0.068)	-0.126	(0.091)
D_{Q_3}	-0.210**	(0.079)	-0.142	(0.103)
D_{Q_4}	-0.369**	(0.103)	-0.433**	(0.130)
$D_{1998\text{-}2000}$	-0.108†	(0.060)	0.561**	(0.080)
	Share of innovative sales (in logit)			
$(\text{R\&D/employee})_{t-1}$ (log)	0.029**	(0.008)	0.030**	(0.011)
$(D_{\text{non-contin. R\&D}})_{t-1}$	-0.028†	(0.016)	-0.042*	(0.017)
Employment (log)	0.499**	(0.053)	0.301**	(0.059)
Technology push	-0.163†	(0.088)	1.024**	(0.278)
Demand pull	0.164†	(0.095)	-0.654**	(0.179)
Distance to frontier				
D_{Q_2}	-0.347**	(0.070)	-0.227**	(0.060)
D_{Q_3}	-0.581**	(0.116)	-0.376**	(0.099)
D_{Q_4}	-0.886**	(0.178)	-0.594**	(0.156)
$D_{1998\text{-}2000}$	-0.608**	(0.117)	1.388**	(0.113)
	Labor productivity, sales/employee (in log)			
Share of innov. sales, latent	1.529**	(0.309)	2.523**	(0.665)
Employment (log)	-0.771**	(0.188)	-0.840**	(0.261)
$D_{1998\text{-}2000}$	0.900**	(0.289)	-3.445**	(0.916)
	Extra parameters			
Individual effects				
σ_η	0.776**	(0.104)	0.947**	(0.116)
σ_{α_3}	2.192**	(0.473)	3.107**	(0.861)
δ_2	1.795**	(0.262)	1.289**	(0.176)
$\rho_{\eta\alpha_3}$	-0.986**	(0.006)	-0.995**	(0.003)
Idiosyncratic errors				
σ_{ϵ_2}	2.320**	(0.081)	1.665**	(0.079)
σ_{ϵ_3}	3.564**	(0.717)	4.241**	(1.103)
ρ_{12}	0.950**	(0.014)	0.834**	(0.033)
ρ_{13}	-0.935**	(0.015)	-0.816**	(0.034)
ρ_{23}	-0.995**	(0.002)	-0.993**	(0.003)
# observations	2645		1910	
Log-likelihood	-6359.598		-4519.029	

‡Three dummies of category of industry and an intercept are included in each equation of the model.
Significance levels : † : 10% * : 5% ** : 1%

The results of the former table are rather disturbing and unexpected. Firstly, the estimated coefficient of the (latent) share of innovative sales in the productivity equation is unexpectedly large compared to what is usually found in the empirical literature. Secondly, the estimated time dummy coefficient has an unexpected sign in the productivity equation for France because Table 5

Table 7: FIML estimates of variant 1 of the panel data CDM model with restrictions on the cross-equation correlations: Unbalanced panel from Dutch and French CIS 2, CIS 3 and CIS 4[‡]

Variable	Coefficient	(Std. Err.)	Coefficient	(Std. Err.)
	France		The Netherlands	
	Innovation incidence, product innovator			
Employment (log)	0.312**	(0.038)	0.304**	(0.048)
Market share (log)	0.049†	(0.026)	0.046	(0.030)
Technology push	0.327	(0.395)	0.197	(0.613)
Demand pull	2.185**	(0.416)	2.057**	(0.573)
Distance to frontier				
D_{Q_2}	0.107	(0.084)	-0.051	(0.103)
D_{Q_3}	-0.106	(0.088)	-0.109	(0.114)
D_{Q_4}	-0.339**	(0.106)	-0.468**	(0.134)
$D_{1998\text{-}2000}$	-0.115†	(0.060)	0.563**	(0.079)
	Share of innovative sales (in logit)			
$(\text{R\&D/employee})_{t\text{-}1}$ (log)	0.072†	(0.041)	0.188**	(0.052)
$(D_{\text{non-contin. R\&D}})_{t\text{-}1}$	-0.246*	(0.104)	-0.201*	(0.102)
Employment (log)	0.447**	(0.054)	0.213**	(0.059)
Technology push	-0.851	(0.731)	1.718†	(0.900)
Demand pull	4.869**	(0.831)	2.379**	(0.839)
Distance to frontier				
D_{Q_2}	0.054	(0.156)	-0.152	(0.147)
D_{Q_3}	-0.428*	(0.168)	-0.445**	(0.162)
D_{Q_4}	-0.979**	(0.205)	-1.022**	(0.197)
$D_{1998\text{-}2000}$	-0.564**	(0.119)	1.475**	(0.114)
	Labor productivity, sales/employee (in log)			
Share of innov. sales, latent	0.021**	(0.005)	0.068**	(0.012)
Employment (log)	0.058**	(0.012)	-0.065**	(0.019)
$D_{1998\text{-}2000}$	-0.196**	(0.020)	-0.182**	(0.035)
	Extra parameters			
Individual effects				
σ_η	0.760**	(0.105)	0.920**	(0.115)
σ_{α_3}	0.578**	(0.013)	0.577**	(0.025)
δ_2	1.649**	(0.278)	1.262**	(0.181)
$\rho_{\eta\alpha_3}$	0, assumed		0, assumed	
Idiosyncratic errors				
σ_{ϵ_2}	2.386**	(0.086)	1.710**	(0.084)
σ_{ϵ_3}	0.398**	(0.010)	0.573**	(0.019)
ρ_{12}	0.965**	(0.019)	0.858**	(0.031)
ρ_{13}	0, assumed		0, assumed	
ρ_{23}	0, assumed		0, assumed	
# observations	2645		1910	
Log-likelihood	-7158.634		-5146.348	

[‡]Three dummies of category of industry and an intercept are included in each equation of the model.
Significance levels : † : 10% ∗ : 5% ∗∗ : 1%

shows that sales per capita increases over time for both countries. Thirdly, the very large negative and significant effect of employment on productivity for both countries seems rather disturbing and is at odds with the finding of most empirical studies.[11] Fourthly, the estimates of technology push and demand pull in the two innovation equations do not tell a consistent story, as their effect goes

[11]See column 'Result/Performance' of Table 1.

both directions (positive or negative) more or less significantly. Finally, it is hard to believe that all the correlations ρ_{13}, $\rho_{\eta\alpha_3}$, and ρ_{23} are negative. The very estimation of these correlations seems problematic. While it is easy to interpret ρ_{12} which gives an estimation of the magnitude and the direction of sample selection bias, it is rather difficult to interpret ρ_{13}, $\rho_{\eta\alpha_3}$, and ρ_{23}. Indeed, since our measure of productivity is observed for both product and non-product innovators, what are we capturing in ρ_{13} and $\rho_{\eta\alpha_3}$? Furthermore, since the (latent) share of innovative sales is already included in the productivity equation, what are we capturing in ρ_{23} and $\rho_{\eta\alpha_3}$? Due to all these problems, we impose the above-mentioned restriction on the cross-equation correlations and estimate the model whose results are reported in Table 7. They are more plausible than those reported in Table 6 and hence will be the basis of our conclusion.

Four-year lagged R&D has a positive and significant effect, albeit less evident for the amount of R&D in the case of France, on the share of innovative sales. This result is more realistic than the instantaneous effect assumed by most empirical studies. The share of innovative sales of both actual and potential product innovators affects instantaneously, positively and significantly labor productivity which is, *ceteris paribus*, smaller during the period 1998-2000 than during the period 2002-2004. The incidence of product innovation increases significantly with firm size (employment) that also affects positively and significantly the share of innovative sales of product innovators. Market share, on the other hand, hardly affects the incidence of product innovation. Thus, Schumpeter's hypothesis is partially supported by our data, i.e., only when size is measured on an absolute basis. Product innovation, both in terms of incidence and intensity, seems to be more market driven than technology pushed. Distance to technological frontier seems to matter in the sense that firms that are closer to the frontier have more important innovation. The result is more pronounced for the intensity of innovation than for its sole incidence. Finally, individual effects have to be accounted for, as shown by the highly significant standard errors σ_η and σ_{α_3}, and sample selection bias must be controlled for, as shown by the large positive and significant estimate of ρ_{12}.

All the results presented so far are similar across countries in terms of significance and direction of the effects of the explanatory variables.[12] However, two main significant differences are observed across the two countries. The first difference is that innovation is, *ceteris paribus,* smaller during the period 1998-200 than during the period 2002-2004 for France, whereas it is the opposite for The Netherlands. This result confirms the pattern observed over time in the proportion of product innovators and in the average share of innovative sales, as shown in Table 5. The second difference

[12] We have not mentioned the significant differences in the magnitude of the effects, e.g. R&D per capita, because we would need to calculate marginal effects to better assess such differences.

is that, *ceteris paribus*, larger firms seem to be more productive in France, whereas smaller firms seem to be so for The Netherlands.

6 Conclusion

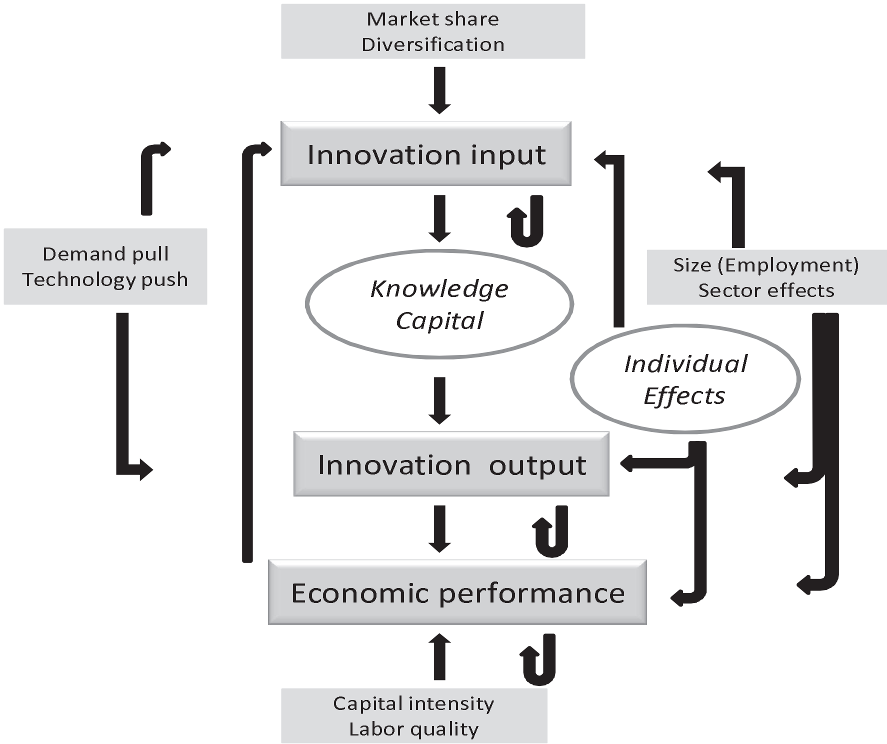
This paper gives, for the first time in the empirical literature, insights into the relationship between R&D, innovation and economic performance using enterprise panel data from three waves of the Dutch and French CIS. A simultaneous-equation model is estimated for the manufacturing sector where (one component of) dynamics and individual effects are accounted for in the relationship, and sample selection bias is controlled for. For both countries, we have the following results. Firstly, four-year lagged R&D has a positive and significant effect, albeit less evident for the amount of R&D in the case of France, on the share of innovative sales. Secondly, the latent share of innovative sales, i.e. for both actual and potential product innovators, affects instantaneously, positively and significantly labor productivity which is, *ceteris paribus*, smaller during the period 1998-2000 than during the period 2002-2004. Thirdly, the incidence of product innovation increases significantly with firm employment that also affects positively and significantly the share of innovative sales of product innovators, while market share hardly affects the incidence of product innovation. Fourthly, product innovation, both in terms of incidence and intensity, seems to be more market driven than technology pushed. Fifthly, distance to technological frontier seems to matter in the sense that firms that are closer to the frontier have more important innovation, which is more pronounced for the intensity of innovation than for its sole incidence. Finally, individual effects have to be accounted for, and sample selection bias must be controlled for. We also observe significant differences across the two countries. The first difference is that innovation is, *ceteris paribus*, smaller during the period 1998-200 than during the period 2002-2004 for France, whereas it is the opposite for The Netherlands. The second difference is that, *ceteris paribus*, larger firms seem to be more productive in France, whereas smaller firms seem to be so for The Netherlands.

In order to cope with the limitations of the analysis and to improve upon it, we plan to achieve the following. Firstly, physical capita is to be included in the augmented production function, which requires merging additional data from other sources. Secondly, we plan to carry out the analysis also for productivity growth. Thirdly, additional dynamics is to be specified in the innovation process, namely persistence of innovation both in terms of incidence and intensity. Fourthly, we could also specify a feedback effect operating from labor productivity to innovation incidence. Fifthly, we could also model R&D activities which is taken predetermined in this analysis. This paper present the results only for the variant of the model with common factor individual effects.

Those of variant 2 of the model are also to be reported for all additional extensions. The final aim of the analysis is to provide a clear-cut answer as to how far we can go in controlling for individual effects, all components of dynamics, sample selection bias and endogeneity of the key variables when estimating the CDM model using panel data from innovation surveys. It should serve as the basis for future work on this topic.

Appendix A CDM framework updated

Figure 1: The CDM framework updated



Appendix B Trivariate normal density

The trivariate normal density function of the individual effects of variant 2 of the model, denoted by $g_3(\alpha_{1i}, \alpha_{2i}, \underline{\underline{\alpha}}_{3i}|...)$, is written as

$$g_3(\alpha_{1i}, \alpha_{2i}, \underline{\underline{\alpha}}_{3i}|...) = \Lambda e^{\frac{-1}{2}\left(A_{11}\frac{\alpha_{1i}^2}{\sigma_{\alpha_1}^2} + 2A_{12}\frac{\alpha_{1i}}{\sigma_{\alpha_1}}\frac{\alpha_{2i}}{\sigma_{\alpha_2}} + 2A_{13}\frac{\alpha_{1i}}{\sigma_{\alpha_1}}\frac{\underline{\underline{\alpha}}_{3i}}{\sigma_{\underline{\underline{\alpha}}_3}} + 2A_{23}\frac{\alpha_{2i}}{\sigma_{\alpha_2}}\frac{\underline{\underline{\alpha}}_{3i}}{\sigma_{\underline{\underline{\alpha}}_3}} + A_{22}\frac{\alpha_{2i}^2}{\sigma_{\alpha_2}^2} + A_{33}\frac{\underline{\underline{\alpha}}_{3i}^2}{\sigma_{\underline{\underline{\alpha}}_3}^2}\right)}, \tag{B.0.11}$$

where

$$\Lambda = \left(\sigma_{\alpha_1}\sigma_{\alpha_2}\sigma_{\underline{\underline{\alpha}}_3}\right)^{-1}(2\pi)^{\frac{-3}{2}}(\Delta)^{\frac{-1}{2}}, \tag{B.0.12}$$

and the expressions of Δ and A_{jk} $(j, k = 1, 2, 3;\ A_{jk} = A_{kj})$ are given by

$$
\begin{aligned}
\Delta &= 1 - \rho^2_{\alpha_1\alpha_2} - \rho^2_{\alpha_1\underline{\underline{\alpha}}_3} - \rho^2_{\alpha_2\underline{\underline{\alpha}}_3} + 2\rho_{\alpha_1\alpha_2}\rho_{\alpha_1\underline{\underline{\alpha}}_3}\rho_{\alpha_2\underline{\underline{\alpha}}_3}, \\
A_{11} &= \frac{1-\rho^2_{\alpha_2\underline{\underline{\alpha}}_3}}{\Delta},\ A_{12} = \frac{\rho_{\alpha_1\underline{\underline{\alpha}}_3}\rho_{\alpha_2\underline{\underline{\alpha}}_3} - \rho_{\alpha_1\alpha_2}}{\Delta}, \\
A_{22} &= \frac{1-\rho^2_{\alpha_1\underline{\underline{\alpha}}_3}}{\Delta},\ A_{23} = \frac{\rho_{\alpha_1\alpha_2}\rho_{\alpha_1\underline{\underline{\alpha}}_3} - \rho_{\alpha_2\underline{\underline{\alpha}}_3}}{\Delta}, \\
A_{33} &= \frac{1-\rho^2_{\alpha_1\alpha_2}}{\Delta},\ A_{13} = \frac{\rho_{\alpha_1\alpha_2}\rho_{\alpha_2\underline{\underline{\alpha}}_3} - \rho_{\alpha_1\underline{\underline{\alpha}}_3}}{\Delta}.
\end{aligned}
\tag{B.0.13}
$$

References

Abramovitz, M. (1956): "Resource and Output Trends in the United States Since 1870," *American Economic Review (Papers and Proceedings)*, 46, 5–23.

Abramowitz, M. and I. Stegun (1964): *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, Washington, D.C..: National Bureau of Standards Applied Mathematics, US Government Printing Office.

Acemoglu, D., P. Aghion, and F. Zilibotti (2003): "Vertical Integration and Distance to Frontier," *Journal of the European Economic Association*, 1, 630–638.

——— (2006): "Distance to Frontier, Selection, and Economic Growth," *Journal of the European Economic Association*, 4, 37–74.

Benavente, J. M. (2006): "The Role of Research and Innovation in Promoting Productivity in Chile," *Economics of Innovation and New Technology*, 15, 301–315.

Bhattacharya, S. and J. R. Ritter (1983): "Innovation and Communication: Signalling with Partial Disclosure," *Review of Economic Studies*, 50, 331–346.

Crépon, B., E. Duguet, and J. Mairesse (1998): "Research, Innovation and Productivity: An Econometric Analysis at the Firm Level," *Economics of Innovation and New Technology*, 7, 115–158.

Griffith, R., E. Huergo, J. Mairesse, and B. Peters (2006): "Innovation and Productivity across four European Countries," *Oxford Review of Economic Policy*, 22, 483–498.

Griliches, Z. (1963): "The Sources of Measured Productivity Growth: United States Agriculture, 1940-1960," *Journal of Political Economy*, 71, 331–346.

——— (1979): "Issues in Assessing the Contribution of Research and Development to Productivity Growth," *Bell Journal of Economics*, 10, 92–116.

——— (1994): "Productivity, R&D and the Data Constraint," *American Economic Review*, 84, 1–23.

Griliches, Z. and J. Mairesse (1984): "Productivity and R&D at the Firm Level," in *R&D, patents and productivity*, ed. by Z. Griliches, Chicago: University of Chicago Press, 339–374.

Jefferson, G. H., B. Huamao, G. Xiaojing, and Y. Xiaoyun (2006): "R&D Performance in Chinese Industry," *Economics of Innovation and New Technology*, 15, 345–366.

Klomp, L. and G. van Leeuwen (2001): "Linking Innovation and Firm Performance: A New Approach," *International Journal of the Economics of Business*, 8, 343–364.

Lööf, H. and A. Heshmati (2002): "Knowledge Capital and Performance Heterogeneity: A Firm-Level Innovation Study," *International Journal of Production Economics*, 76, 61–85.

——— (2006): "On the Relationship Between Innovation and Performance: A Sensitivity Analysis," *Economics of Innovation and New Technology*, 15, 317–344.

OECD (2005): *Oslo Manual, Guidelines for Collecting and Interpreting Innovation Data*, Paris: OECD Publishing, 3rd ed.

——— (2007): *OECD Science, Technology and Industry Scoreboard 2007: Innovation and Performance in the Global Economy*, Paris: OECD; 2007 edition.

Pakes, A. and Z. Griliches (1980): "Patents and R&D at the Firm Level: A First Report," *Economics Letters*, 5, 377–381.

Parisi, M. L., F. Schiantarelli, and A. Sembenelli (2006): "Productivity, Innovation and R&D: Micro Evidence for Italy," *European Economic Review*, 50, 2037–2061.

Scherer, F. M. (1965): "Firm Size, Market Structure, Opportunity and the Output of Patented Inventions," *American Economic Review*, 55, 1097–1125.

Schmookler, J. A. (1966): *Invention and Economic Growth*, Cambridge: Harvard University Press.

Schumpeter, J. A. (1934): *The Theory of Economic Development*, Cambridge, MA: Harvard University Press.

——— (1942): *Capitalism, Socialism and Democracy*, New York: Harper and Brothers.

Solow, R. M. (1957): "Technical Change and the Aggregate Production Function," *Review of Economics and Statistics*, 39, 312–320.

van Leeuwen, G. (2002): "Linking Innovation to Productivity Growth Using two Waves of the Community Innovation Survey," OECD Science, Technology and Industry Working Papers, 2002/8, OECD Publishing.

van Leeuwen, G. (2002): "Linking Innovation to Productivity Growth Using two Waves of the Community Innovation Survey," OECD Science, Technology and Industry Working Papers, 2002/8, OECD Publishing.

van Leeuwen, G. and L. Klomp (2006): "On the Contribution of Innovation to Multi-Factor Productivity Growth," *Economics of Innovation and New Technology*, 15, 367–390.

Vella, F. (1998): "Estimating Models with Sample Selection Bias: A Survey," *Journal of Human Resources*, 33, 127–169.