Innovative Sales, R&D and Other Innovation Expenditures: Are there Lags? Estimating from Panel Data Vector Autoregressions

Wladimir Raymond, Pierre Mohnen, Franz Palm and Sybrand Schim van der Loeff[1]

March 11, 2008

Abstract

This paper studies the dynamic relationship between innovation input and output in Dutch manufacturing using an unbalanced panel of enterprise data from four waves of the Community Innovation Survey between 1994-2002. We estimate by maximum likelihood a panel data vector autoregression accounting for individual effects. There is strong evidence of a lag and feedback effect only between R&D and the share of innovative sales in the high-tech sector. We find persistence of innovation input and output, and simultaneity between them. The two types of innovation inputs are intertemporal substitutes. Individual effects play an important role in the relationship.

[1]The empirical part of this study has been carried out using the remote access facility of the Centre for Research of Economic Microdata of Statistics Netherlands. The authors wish to thank Statistics Netherlands for helping them in accessing and using the Micronoom data set. The views expressed in this paper are solely those of the authors. The first author acknowledges financial support from METEOR. The first and third author acknowledge financial support by the Royal Netherlands Academy of Arts and Sciences. Please address correspondence to: Wladimir Raymond, University of Maastricht, P.O. Box 616, 6200 MD Maastricht, The Netherlands; Tel.: (+31) 43 388 4962; Fax: (+31) 43 388 4874; w.raymond@ke.unimaas.nl.

1 Introduction

This paper studies the dynamic relationship between innovation input and innovation output in Dutch manufacturing using an unbalanced panel of enterprise data from four waves of the Community Innovation Survey (CIS) pertaining to the periods 1994-1996, 1996-1998, 1998-2000 and 2000-2002.

The first attempt to study such a dynamic relationship is by Pakes and Griliches (1980a,b) who define a knowledge production function (KPF) relating innovation input proxied by R&D expenditures to knowledge increment proxied by patents. They estimate a distributed lag (log-log) regression where patents are regressed on current and 5 lagged R&D variables. Accounting for fixed-effects that are correlated with R&D, they find a positive and significant effect of current (simultaneity) and 5-year lagged R&D on patents (lag truncation), and nothing significant in between.[2] Hausman et al. (1984) and Hall et al. (1986) estimate several distributed lag specifications of the patents-R&D relationship, namely log-log, Poisson and negative binomial regressions with absence and presence of individual effects. In their preferred specifications,[3] they find as main results no evidence of a lag effect of R&D on patents but only simultaneity between them.

The purpose of the paper is to estimate a KPF using another indicator of innovation output than patents, namely the share of innovative sales, and two indicators of innovation input, namely R&D and non-R&D innovation expenditures (e.g. purchase of rights and licenses). More specifically, we estimate two specifications of the KPF. In one specification, a single innovation input enters the KPF where R&D expenditures and other non-R&D innovation expenditures interchange as innovation input, while in the second one both innovation inputs enter together the KPF. In addition to the time lag between innovation input and innovation output modeled in the previously-mentioned studies, we allow for persistence in innovation input and innovation output intensity, and a feedback effect of innovation output on innovation input. The simultaneity between innovation input and innovation output that is identified in the above-mentioned studies is explicitly accounted for by allowing for cross-equation correlations between the disturbance terms. Finally, like the above-mentioned studies, individual effects possibly correlated with the regressors are taken into account. The resulting model is a panel data vector autoregression (VAR) with a two-component vector of dependent variables in the single input specification and a three-component vector in the double input specification.

This study contributes to the empirical literature in various ways. First, a panel data knowledge

[2]The lag truncation effect means that when allowing for longer lags in R&D (say 6- and 7-year lags) the most recent lagged R&D variables (say 5-year lag) are no longer significant.

[3]The preferred specifications are those where the discreteness of patents and the presence of individual effects correlated with R&D are accounted for.

production function with two innovation inputs is estimated for the first time using several waves of the Dutch CIS. Secondly, the KPF includes various features all of which are not accounted for in a unified framework in the empirical literature. Thirdly, we propose an alternative to the instrumental variable quasi-differenced method of Holtz-Eakin et al. (1988) that cannot be applied to data that have similar characteristics to the CIS, namely those of a short period ($\max(T_i) = 4$) panel with many qualitative variables and a few continuous ones that show little within variation. We estimate the model by maximum likelihood using multiple step Gauss-Hermite quadrature, and Wooldridge's (2005) solution to handle the initial conditions problem. The results suggest strong persistence of innovation input, regardless of how it is measured, and of innovation output. There is a lag effect between innovation input and innovation output that differs according to the type of innovation input and the type of sector. Similarly, the feedback effect of innovation output on innovation input differs according to the type of innovation input and the type of the sector. Furthermore, there is evidence of simultaneity between innovation input and innovation output, regardless of how innovation input is measured, and intertemporal substitutability between both types of innovation input. Finally, unobserved heterogeneity plays an important role, through individual effects, in the KPF.

The remainder of the paper is organized as follows. Section 2 describes the data, Section 3 presents the two specifications of the KPF that are estimated in Section 4. The estimation results are discussed in Section 5, and we summarize and conclude in Section 6.

2 Data

The data are collected by the *Centraal Bureau voor de Statistiek* (CBS) and stem from four waves of the Dutch CIS, namely CIS 2 (1994-1996), CIS 2.5 (1996-1998), CIS 3 (1998-2000) and CIS 3.5 (2000-2002), merged with data from the Production Survey (PS). Only enterprises in Dutch manufacturing (SBI 15-37) are included in the analysis.[4] The population of interest consists of enterprises with at least ten employees and positive sales at the end of each period covered by the innovation survey.

The CIS and PS data are collected at the enterprise level. A combination of a census and a stratified random sampling is used for each wave of the CIS and PS. A census is used for the population of enterprises with at least 50 employees, and a stratified random sampling is used for enterprises with less than 50 employees. 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

[4]SBI stands for the Dutch standard industrial classification and gives the enterprise economic activity.

wave of the CIS and PS resulting in about 3000 enterprises in each wave of the merged data of our sample.

Table 1: Patterns of the merged data

No.	Pattern	Frequency	Percent
1	0001	1013	14.677
2	1000	969	14.039
3	0100	870	12.605
4	0010	851	12.330
5	1100	754	10.924
6	1111	588	8.519
7	0011	346	5.013
8	1110	285	4.129
9	1101	257	3.724
10	0111	218	3.159
11	0110	183	2.651
12	0101	173	2.507
13	1001	136	1.970
14	1011	133	1.927
15	1010	126	1.823
		6902	100.000

Table 1 shows the patterns of the merged data. For instance, 1013 enterprises, i.e. 15% of the merged data, are present only in the fourth wave of the CIS and PS. The sample of analysis consists of enterprises that take part in at least two consecutive innovation and production surveys (patterns 5-11 and 14) resulting in an unbalanced panel of 2764 enterprises. Thus, we have in our sample three categories of enterprises. The first category consists of new enterprises and enterprises that were not sampled in, or responded to, at least one previous innovation or production survey. The second category includes enterprises that died, were merged or acquired, or ceased to be sampled or failed to respond after two or three consecutive innovation or production waves. The last category mainly consists of large firms that existed in 1994, survived without merger and acquisition until 2002 and took part in all four waves of the innovation and production surveys, hence forming a balanced panel of 588 enterprises (pattern 6). The sample as it stands now includes innovative and non-innovative enterprises. An innovative enterprise is defined as one that has positive total innovation expenditures at the end of the period of the survey.[5] We further restrict the analysis sample by considering enterprises that are innovative in two consecutive waves of the CIS, hence resulting in a conditional analysis.[6]

[5]In addition to R&D, innovation expenditures consist of expenditures for technical preparations to realize the actual implementation of product and process innovations, purchase of rights and licenses to use external technology, expenditures for marketing activities aimed at market introduction of product innovations, and expenditures for staff training aimed at the development and/or introduction of new product and process innovations. Hence, an innovative enterprise may have only R&D activities, may have no R&D activities but other innovation activities, or may have both R&D and other innovation activities.

[6]The reason for a conditional analysis is that the regressors of the VAR models are only available for innovative enterprises.

2.1 Dependent variables

As we control for simultaneity between innovation inputs and innovation output, we distinguish three dependent variables.

The first dependent variable is R&D intensity calculated as the ratio of total (intramural and extramural) R&D expenditures from the CIS over total sales stemming from the PS. This variable is measured at the end of the period under review. Its logarithmic transformation is used in the estimation. Since an innovative enterprise does not necessarily have R&D activities, we use the pragmatic solution of substituting a small but positive value ε for the zero values of R&D intensity.

The second dependent variable is (other) innovation input intensity calculated as the ratio of innovation expenditures (excluding R&D) also stemming from the CIS over total sales. Its logarithmic transformation is also used in the estimation, and a solution similar to that of R&D is applied when innovation input intensity takes on the value 0.

The CIS data set also provides information regarding the share in total sales accounted for by sales of new or improved products, measured at the end of the period under review. This is the measure of innovation output intensity used in this study. A logit transformation of this measure is used in order to make it lie within the set of real numbers.[7]

2.2 Explanatory variables

Besides simultaneity in the knowledge production function, we account for dynamics in the innovation process. More specifically, we study the persistence of innovation inputs and that of innovation output, the lag between innovation inputs and innovation output, and the feedback effects of innovation output on innovation inputs. It is to be noted that a one-period lag actually corresponds to two years.

Hence, we explain current R&D intensity, current innovation input intensity and current share of innovative sales by their own lagged values (persistence), cross-lagged values (time lag and feedback effects), and lagged values of other strictly exogenous explanatory variables. Furthermore, we carry out the analysis separately for high-tech and low-tech sectors as defined by OECD (1999),[8] and include in each equation and for each sector industry (according to 2-digit SBI) and time dummies.[9] The set of strictly exogenous explanatory variables includes size stemming from the PS, and indicators for technology push, demand pull and subsidy stemming from the CIS. Size

[7]The share of innovative sales takes on the values 0 for process-only innovators, and 1 for innovators that are newly established. They are replaced respectively by 0.001 and 0.99 in the logit transformation.

[8]The high-tech sector, as defined by OECD (1999), consists of the industries of chemicals, electrical products, machinery and equipment and vehicles. The low-tech sector consists of the industries of food, metals, non-metallic products, plastics, products not elsewhere classified, textiles and wood.

[9]The economic activity of an enterprise is defined by CBS up to a 5-digit SBI level. However, because of confidentiality reasons, we use a 2-digit level classification of economic activity in this study.

and technology push are included on the grounds of the Schumpeterian tradition, demand pull is included according to Schmookler, and subsidized enterprises are expected to be more innovative although evidence on this score is mixed (David et al., 2000). Size is measured by the number of employees at the end of the period under survey and is log-transformed in the estimation. Technology push is proxied by a dummy variable constructed from the indicators stating the importance of public or private research institutions (e.g. universities) as sources of information for innovation. This proxy takes on the value one if at least one of these institutions is deemed to be important or very important to an enterprise (i.e. has value 2 or 3 on a 0-3 Likert scale), and zero otherwise. To proxy demand pull, we construct a dummy variable that equals one if at least one of the following objectives of innovation is given the highest mark on a 0-3 Likert scale, and zero otherwise: "open-up new markets", "extend product range" and "replace products phased out". If an enterprise answers that it has been granted at least one kind of subsidy during the period under review, the variable subsidy takes on the value one and zero otherwise.

2.3 Descriptive statistics

Table 2 reports descriptive statistics of the dependent and explanatory variables for both sectors, as well as t- and z-test results of equality of means and percentages across sectors. The table suggests that enterprises that belong to the high-tech sector spend on average significantly larger amounts on R&D than the low-tech counterparts. However, when R&D expenditures are scaled by total sales there is no strong evidence against the null hypothesis of equality of R&D across sectors. The reason is that enterprises are equally large on average across sectors (see row 6), and that R&D intensity implicitly accounts for size since this variable can also be proxied by total sales.[10] As for non-R&D innovation expenditures, they are not statistically and significantly different across sectors, whether scaled by total sales or not. The remaining variables, namely the share of innovative sales, technology push, demand pull and subsidy are statistically and significantly (at 1% level) different across sectors.

3 Knowledge production functions

Two specifications of the knowledge production function are considered. In one specification innovation output, i.e. the share of innovative sales, is determined by a single innovation input, and in the other one innovation output is determined by two innovation inputs. In the single input

[10] Another reason why R&D intensity is not statistically and significantly different on average across sectors originates from the conditional feature of our analysis. When performing a similar test of equality of means of R&D intensity and size using a larger sample that includes enterprises that are not necessarily innovative in two consecutive waves, the test statistics are significant at 1% level.

Table 2: Descriptive statistics and tests of equality of means and percentages across sectors

Variable	Mean	Std. Dev. Overall	Std. Dev. Between	Std. Dev. Within	Mean	Std. Dev. Overall	Std. Dev. Between	Std. Dev. Within	Test of equality of mean and percentage
	High-tech				Low-tech				
R&D expenditures	4855.057	25031.030	18042.310	7121.499	675.774	5029.986	3526.284	1559.282	|t|=5.968; p=0.000
R&D intensity†	0.259	3.149	2.990	1.817	0.108	1.088	1.150	0.466	|t|=1.633; p=0.103
Other innov. expenditures	529.301	2532.793	1832.360	1443.154	426.633	1860.679	1425.734	1031.350	|t|=1.138; p=0.255
Other innov. intensity‡	0.054	0.729	0.857	0.235	0.222	4.381	4.252	2.310	|t|=1.217; p=0.224
Share of innovative sales††	0.329	0.272	0.254	0.123	0.230	0.219	0.211	0.097	|t|=9.874; p=0.000
Size‡‡	289.398	788.399	507.838	454.577	249.836				|t|=1.005; p=0.315
Technology push	0.278				0.213				|z|=3.537; p=0.000
Demand pull	0.707				0.651				|z|=2.935; p=0.003
Subsidy	0.585				0.413				|z|=8.291; p=0.000

† ln(R&D/total sales); ‡ ln(other innovation expenditures/total sales); †† logit(innovative sales/ total sales); ‡‡ ln(number of employees) are used in the estimation.

specification R&D and other innovation intensity interchange as innovation input, while they enter together the knowledge production function in the double input specification. Both specifications account for simultaneity between innovation inputs and innovation output, dynamics in the innovation process and unobserved heterogeneity through individual effects. Hence, the empirical model is a panel data vector autoregression with a two-component vector of dependent variables in the single input specification, and a three-component vector of dependent variables in the double input specification.

Single input KPF

The single input KPF is written as

$$IN_{it} = \gamma_{11} IN_{i,t-1} + \gamma_{12} OUT_{i,t-1} + \boldsymbol{\beta}_1' \mathbf{x}_{1i,t-1} + \mu_{1it}, \tag{1}$$

$$OUT_{it} = \gamma_{21} IN_{i,t-1} + \gamma_{22} OUT_{i,t-1} + \boldsymbol{\beta}_2' \mathbf{x}_{2i,t-1} + \mu_{2it}, \tag{2}$$

where $i = 1, ...N$, $t = 1, ...T_i$, IN denotes innovation input conducive to innovation output OUT, $\mathbf{x}_j$ $(j = 1, 2)$ denotes vectors of strictly exogenous explanatory variables,[11] and μ_j denotes the error terms capturing unobserved variables that affect innovation input and innovation output, γ_{jk} and $\boldsymbol{\beta}_j'$ $(j, k = 1, 2)$ are parameters to be estimated. γ_{11} and γ_{22} capture respectively the persistence of the intensity of innovation input and innovation output, γ_{21} captures the lag effect between innovation input and innovation output, and γ_{12} captures the feedback effect of innovation output on innovation input.

As mentioned earlier, two versions of the single input KPF are estimated. Innovation input (eq. (1)) is proxied by R&D intensity in one version and by other innovation intensity in the other one, and is estimated together with innovation output (eq. (2)) proxied by the share of innovative sales.

Double input KPF

Similarly, the double input KPF is written as

$$IN_{1it} = \gamma_{11} IN_{1i,t-1} + \gamma_{12} IN_{2i,t-1} + \gamma_{13} OUT_{i,t-1} + \boldsymbol{\beta}_1' \mathbf{x}_{1i,t-1} + \mu_{1it}, \tag{3}$$

$$IN_{2it} = \gamma_{21} IN_{1i,t-1} + \gamma_{22} IN_{2i,t-1} + \gamma_{23} OUT_{i,t-1} + \boldsymbol{\beta}_2' \mathbf{x}_{2i,t-1} + \mu_{2it}, \tag{4}$$

$$OUT_{it} = \gamma_{31} IN_{1i,t-1} + \gamma_{32} IN_{2i,t-1} + \gamma_{33} OUT_{i,t-1} + \boldsymbol{\beta}_3' \mathbf{x}_{3i,t-1} + \mu_{3it}, \tag{5}$$

[11] The explanatory variables included in $\mathbf{x}_1$ and $\mathbf{x}_2$ are similar in this study although they need not be.

where both innovation inputs, R&D intensity and other innovation intensity, enter together the knowledge production function. In this specification, the persistence of each innovation input, captured by γ_{11} and γ_{22}, is estimated jointly with that of innovation output captured by γ_{33}. Furthermore, we estimate jointly the lag effect between each innovation input and innovation output captured by γ_{31} and γ_{32}, the lag effect of each innovation input on the other captured by γ_{12} and γ_{21}, and the feedback effect of innovation output on each innovation input captured by γ_{13} and γ_{23}. $\mathbf{x}_j$ $(j = 1, 2, 3)$ and μ_j denote respectively the vectors of strictly exogenous variables and unobserved variables that affect innovation inputs and output.[12]

4 Estimation

We estimate the knowledge production functions by maximum likelihood taking account of the features of the CIS data. They are those of a short period ($\max(T_i) = 4$) panel with many qualitative variables (e.g. industry dummies) and a few continuous ones that show little within variation (e.g. size).

The first step to estimation is to model the error structure of the knowledge production functions. Ideally, one should analyze the autocorrelation and partial autocorrelation functions of the error terms as suggested by MaCurdy (1981, 1982). Instead, we start by assuming we assume the simplest form of autocorrelation of the error terms, namely a variance components scheme with stationary individual effects and independently and identically distributed (across individuals and over time) idiosyncratic terms.[13] Formally, the error terms are written as

$$\boldsymbol{\mu}_{it} = \boldsymbol{\alpha}_i + \boldsymbol{\epsilon}_{it}, \tag{6}$$

where $\boldsymbol{\alpha}_i$ and $\boldsymbol{\epsilon}_{it}$ denote respectively the individual effects and the idiosyncratic errors with two components in the single input KPF and three components in the double input KPF. Because of the autoregressive feature of the model, $cor[IN_{is}\mu_{1it}] \neq 0$ and $cor[OUT_{is}\mu_{2it}] \neq 0$ $(s < t)$ in the single input KPF, and $cor[IN_{1is}\mu_{1it}] \neq 0$, $cor[IN_{2is}\mu_{2it}] \neq 0$ and $cor[OUT_{is}\mu_{3it}] \neq 0$ in the double input KPF. These correlations operate through the individual effects. In order to avoid overestimating the persistence parameters, they must be accounted for in each equation of the VAR model. This problem is known in the econometric literature as the "the initial conditions problem". We adopt the Wooldridge (2005) approach of handling the initial conditions problem,

[12] For simplicity, we use the same notation in equations (1) and (2) as in equations (3)-(5). Obviously, the parameters of the single input KPF are different from those of the double input KPF.

[13] As a matter of fact, we have carried out the autocorrelation and partial autocorrelation analysis that is to be included in a later version of the paper together with the resulting estimation results.

i.e. we write the individual effects as

$$\alpha_{1i} = b_{10} + b_{11} IN_{i0} + \mathbf{b}'_{12}\mathbf{x}_{1i} + a_{1i}, \tag{7}$$

$$\alpha_{2i} = b_{20} + b_{21} OUT_{i0} + \mathbf{b}'_{22}\mathbf{x}_{1i} + a_{2i}, \tag{8}$$

when the knowledge production function consists of a single input and

$$\alpha_{1i} = b_{10} + b_{11} IN_{1i0} + \mathbf{b}'_{12}\mathbf{x}_{1i} + a_{1i}, \tag{9}$$

$$\alpha_{2i} = b_{20} + b_{21} IN_{2i0} + \mathbf{b}'_{22}\mathbf{x}_{2i} + a_{2i}, \tag{10}$$

$$\alpha_{3i} = b_{30} + b_{31} OUT_{i0} + \mathbf{b}'_{32}\mathbf{x}_{3i} + a_{3i}, \tag{11}$$

when the knowledge production function consists of two inputs. The additional b-parameters of equations (7)-(8) and (9)-(11) capture not only the "correlation" between the individual effects and the initial conditions but also the correlation between the individual effects and the strictly exogenous explanatory variables, and are to be estimated.[14] The idiosyncratic errors $\boldsymbol{\epsilon}_{it}$ and the individual effects components $\boldsymbol{a}_i$ are assumed to be mutually independent and distributed according to a normal distribution with mean zero and 2×2 covariance matrices in the single input KPF, and 3×3 covariance matrices in the double input KPF. The covariance matrices are written respectively

$$\Sigma_{\boldsymbol{\epsilon}} = \begin{pmatrix} \sigma^2_{\epsilon_1} & \\ \rho_{\epsilon_1\epsilon_2}\sigma_{\epsilon_1}\sigma_{\epsilon_2} & \sigma^2_{\epsilon_2} \end{pmatrix}, \; \Sigma_{\boldsymbol{a}} = \begin{pmatrix} \sigma^2_{a_1} & \\ \rho_{a_1a_2}\sigma_{a_1}\sigma_{a_2} & \sigma^2_{a_2} \end{pmatrix} \tag{12}$$

in the single input KPF and

$$\Sigma_{\boldsymbol{\epsilon}} = \begin{pmatrix} \sigma^2_{\epsilon_1} & & \\ \rho_{\epsilon_1\epsilon_2}\sigma_{\epsilon_1}\sigma_{\epsilon_2} & \sigma^2_{\epsilon_2} & \\ \rho_{\epsilon_1\epsilon_3}\sigma_{\epsilon_1}\sigma_{\epsilon_3} & \rho_{\epsilon_2\epsilon_3}\sigma_{\epsilon_2}\sigma_{\epsilon_3} & \sigma^2_{\epsilon_3} \end{pmatrix}, \; \Sigma_{\boldsymbol{a}} = \begin{pmatrix} \sigma^2_{a_1} & & \\ \rho_{a_1a_2}\sigma_{a_1}\sigma_{a_2} & \sigma^2_{a_2} & \\ \rho_{a_1a_3}\sigma_{a_1}\sigma_{a_3} & \rho_{a_2a_3}\sigma_{a_2}\sigma_{a_3} & \sigma^2_{a_3} \end{pmatrix} \tag{13}$$

in the double input KPF. The parameters of the covariance matrices are also to be estimated.

[14] In order for both the $\mathbf{b}$- and $\boldsymbol{\beta}$-vectors of parameters to be estimated, the strictly exogenous explanatory variables must exhibit sufficient within variation. This is hardly the case in our data. Hence, we can control only for the correlation between the individual effects and the initial conditions.

Define

$$
\begin{aligned}
A_{1it} &= \gamma_{11} IN_{i,t-1} + \gamma_{12} OUT_{i,t-1} + \boldsymbol{\beta}_1' \mathbf{x}_{1i,t-1} + b_{10} + b_{11} IN_{i0} + \mathbf{b}_{12}' \mathbf{x}_{1i}, \\
A_{2it} &= \gamma_{21} IN_{i,t-1} + \gamma_{22} OUT_{i,t-1} + \boldsymbol{\beta}_2' \mathbf{x}_{2i,t-1} + b_{20} + b_{21} OUT_{i0} + \mathbf{b}_{22}' \mathbf{x}_{2i},
\end{aligned} \tag{14}
$$

and

$$
\begin{aligned}
B_{1it} &= \gamma_{11} IN_{1i,t-1} + \gamma_{12} IN_{2i,t-1} + \gamma_{13} OUT_{i,t-1} + \boldsymbol{\beta}_1' \mathbf{x}_{1i,t-1} + b_{10} + b_{11} IN_{1i0} + \mathbf{b}_{12}' \mathbf{x}_{1i}, \\
B_{2it} &= \gamma_{21} IN_{1i,t-1} + \gamma_{22} IN_{2i,t-1} + \gamma_{23} OUT_{i,t-1} + \boldsymbol{\beta}_2' \mathbf{x}_{2i,t-1} + b_{20} + b_{21} IN_{2i0} + \mathbf{b}_{22}' \mathbf{x}_{2i}, \\
B_{3it} &= \gamma_{31} IN_{1i,t-1} + \gamma_{32} IN_{2i,t-1} + \gamma_{33} OUT_{i,t-1} + \boldsymbol{\beta}_3' \mathbf{x}_{3i,t-1} + b_{30} + b_{31} OUT_{i0} + \mathbf{b}_{32}' \mathbf{x},
\end{aligned} \tag{15}
$$

the individual likelihood function of the single input KPF conditional on the regressors, the initial conditions and the individual effects denoted by $\prod_{t=1}^{T_i} L_{it}(...|..., a_{1i}, a_{2i})$ is written as

$$
\begin{aligned}
&\prod_{t=1}^{T_i} \frac{1}{\sigma_{\epsilon_1}\sqrt{1-\rho^2_{\epsilon_1\epsilon_2}}} \phi\left(\frac{IN_{it} - A_{1it} - a_{1i} - \rho_{\epsilon_1\epsilon_2}\frac{\sigma_{\epsilon_1}}{\sigma_{\epsilon_2}}(OUT_{it} - A_{2it} - a_{2i})}{\sigma_{\epsilon_1}\sqrt{1-\rho^2_{\epsilon_1\epsilon_2}}} \right) \\
&\qquad \times \frac{1}{\sigma_{\epsilon_2}} \phi\left(\frac{OUT_{it} - A_{2it} - a_{2i}}{\sigma_{\epsilon_2}} \right),
\end{aligned} \tag{16}
$$

and that of the double input KPF denoted by $\prod_{t=1}^{T_i} L_{it}(...|..., a_{1i}, a_{2i}, a_{3i})$ is written as

$$
\begin{aligned}
&\prod_{t=1}^{T_i} \frac{1}{\sigma_{\epsilon_1}\sqrt{1-R^2_{1.23}}} \phi\left(\frac{IN_{1it} - B_{1it} - a_{1i} - \rho_{12.3}\frac{\sigma_{\epsilon_1}}{\sigma_{\epsilon_2}}(IN_{2it} - B_{2it} - a_{2i}) - \rho_{13.2}\frac{\sigma_{\epsilon_1}}{\sigma_{\epsilon_3}}(OUT_{it} - B_{3it} - a_{3i})}{\sigma_{\epsilon_1}\sqrt{1-R^2_{1.23}}} \right) \\
&\qquad \times \frac{1}{\sigma_{\epsilon_2}\sqrt{1-\rho^2_{\epsilon_2\epsilon_3}}} \phi\left(\frac{IN_{2it} - B_{2it} - a_{2i} - \rho_{\epsilon_2\epsilon_3}\frac{\sigma_{\epsilon_2}}{\sigma_{\epsilon_3}}(OUT_{it} - B_{3it} - a_{3i})}{\sigma_{\epsilon_2}\sqrt{1-\rho^2_{\epsilon_2\epsilon_3}}} \right) \\
&\qquad\qquad \times \frac{1}{\sigma_{\epsilon_3}} \phi\left(\frac{OUT_{it} - B_{3it} - a_{3i}}{\sigma_{\epsilon_3}} \right),
\end{aligned} \tag{17}
$$

where $R^2_{1.23}$ denotes the multiple correlation of ϵ_{1it} with ϵ_{2it} and ϵ_{3it}, $\rho_{12.3}$ is the partial correlation between ϵ_{1it} and ϵ_{2it} given ϵ_{3it}, $\rho_{13.2}$ is the partial correlation between ϵ_{1it} and ϵ_{3it} given ϵ_{2it}, and ϕ is the univariate standard normal density function. The expressions of the multiple and partial correlations are given by

$$
R^2_{1.23} = \frac{\rho^2_{\epsilon_1\epsilon_2} + \rho^2_{\epsilon_1\epsilon_3} - 2\rho_{\epsilon_1\epsilon_2}\rho_{\epsilon_1\epsilon_3}\rho_{\epsilon_2\epsilon_3}}{1-\rho^2_{\epsilon_2\epsilon_3}}, \; \rho_{12.3} = \frac{\rho_{\epsilon_1\epsilon_2} - \rho_{\epsilon_1\epsilon_3}\rho_{\epsilon_2\epsilon_3}}{1-\rho^2_{\epsilon_2\epsilon_3}} \text{ and } \rho_{13.2} = \frac{\rho_{\epsilon_1\epsilon_3} - \rho_{\epsilon_1\epsilon_2}\rho_{\epsilon_2\epsilon_3}}{1-\rho^2_{\epsilon_2\epsilon_3}}. \tag{18}
$$

The unconditional (with respect to the individual effects) individual likelihood functions are obtained by "integrating out" the individual effects in equations (16) and (17) with respect to their joint density function. They are written as

$$L_i = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\prod_{t=1}^{T_i} L_{it}(...|..., a_{1i}, a_{2i})g(a_{1i}, a_{2i})da_{1i}da_{2i}, \tag{19}$$

in the single input KPF and

$$L_i = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\prod_{t=1}^{T_i} L_{it}(...|..., a_{1i}, a_{2i}, a_{3i})h(a_{1i}, a_{2i}, a_{3i})da_{1i}da_{2i}da_{3i} \tag{20}$$

in the double input KPF, where $g()$ and $h()$ denote respectively the bivariate and trivariate normal distribution. The multiple integrals in equations (19) and (20) are calculated using respectively two- and three-step Gauss-Hermite quadrature which states that

$$\int_{-\infty}^{\infty} e^{-z^2} f(z)dz \simeq \sum_{m=1}^{M} w_m f(a_m), \tag{21}$$

where w_m and a_m are respectively the weights and abscissas of the Gauss-Hermite integration,[15] the tables of which are formulated in mathematical textbooks (e.g. Abramovitz and Stegun, 1964), and M is the total number of integration points. The larger M, the more accurate the Gauss-Hermite approximation. The expressions of the unconditional individual likelihood functions are shown to be

$$\begin{aligned} L_i \simeq & \frac{\sqrt{1-\rho^2_{a_1a_2}}}{\pi}\sum_{p=1}^{P} w_p \left\{\prod_{t=1}^{T_i}\left[\frac{1}{\sigma_{\epsilon_2}}\phi\left(\frac{OUT_{it}-G_{2it}}{\sigma_{\epsilon_2}}\right)\right]\left\{\sum_{m=1}^{M} w_m e^{2\rho_{a_1a_2}a_ma_p}\right.\right. \\ & \left.\left.\prod_{t=1}^{T_i}\frac{1}{\sigma_{\epsilon_1}\sqrt{1-\rho^2_{\epsilon_1\epsilon_2}}}\phi\left(\frac{IN_{it}-G_{1it}-\rho_{\epsilon_1\epsilon_2}\frac{\sigma_{\epsilon_1}}{\sigma_{\epsilon_2}}(OUT_{it}-G_{2it})}{\sigma_{\epsilon_1}\sqrt{1-\rho^2_{\epsilon_1\epsilon_2}}}\right)\right\}\right\}, \end{aligned} \tag{22}$$

in the single input KPF and

$$\begin{aligned} L_i \simeq & \Lambda\sum_{q=1}^{Q} w_q \left\{\prod_{t=1}^{T_i}\left[\frac{1}{\sigma_{\epsilon_3}}\phi\left(\frac{OUT_{it}-H_{3it}}{\sigma_{\epsilon_3}}\right)\right]\sum_{p=1}^{P} w_p \left\{e^{-\frac{2a_pa_q\Delta_{23}}{\sqrt{\Delta_{22}}\sqrt{\Delta_{33}}}}\right.\right. \\ & \prod_{t=1}^{T_i}\left[\frac{1}{\sigma_{\epsilon_2}\sqrt{1-\rho^2_{\epsilon_2\epsilon_3}}}\phi\left(\frac{IN_{2it}-H_{2it}-\rho_{\epsilon_2\epsilon_3}\frac{\sigma_{\epsilon_2}}{\sigma_{\epsilon_3}}(OUT_{it}-H_{3it})}{\sigma_{\epsilon_2}\sqrt{1-\rho^2_{\epsilon_2\epsilon_3}}}\right)\right]\sum_{m=1}^{M} w_m \left\{e^{-\frac{2a_ma_p\Delta_{12}}{\sqrt{\Delta_{11}}\sqrt{\Delta_{22}}}-\frac{2a_ma_q\Delta_{13}}{\sqrt{\Delta_{11}}\sqrt{\Delta_{33}}}}\right. \\ & \left.\left.\left.\prod_{t=1}^{T_i}\left[\frac{1}{\sigma_{\epsilon_1}\sqrt{1-R^2_{1.23}}}\phi\left(\frac{IN_{1it}-H_{1it}-\rho_{12.3}\frac{\sigma_{\epsilon_1}}{\sigma_{\epsilon_2}}(IN_{2it}-H_{2it})-\rho_{13.2}\frac{\sigma_{\epsilon_1}}{\sigma_{\epsilon_3}}(OUT_{it}-H_{3it})}{\sigma_{\epsilon_1}\sqrt{1-R^2_{1.23}}}\right)\right]\right\}\right\}\right\} \end{aligned} \tag{23}$$

in the double input KPF, where w_m, w_p and w_q are the weights of the first-, second- and third-step

[15] In fact the individual effects need not be distributed according to the normal distribution. The method works with any distribution that belongs to the exponential family.

Gauss-Hermite quadrature; a_m, a_p and a_q are the corresponding abscissas, and M, P and Q are the first-, second- and third-step total number of integration points (see Raymond, 2007, chap. 3 and 6). The expressions of $\Delta' s$, Λ, $G' s$ and $H' s$ are given by

$$\Delta = 1 - \rho^2_{a_1a_2} - \rho^2_{a_1a_3} - \rho^2_{a_2a_3} + 2\rho_{a_1a_2}\rho_{a_1a_3}\rho_{a_2a_3},$$

$$\begin{aligned}
\Delta_{11} &= \frac{1-\rho^2_{a_2a_3}}{\Delta},\ \Delta_{22} = \frac{1-\rho^2_{a_1a_3}}{\Delta},\ \Delta_{33} = \frac{1-\rho^2_{a_1a_2}}{\Delta},\\
\Delta_{12} &= \frac{\rho_{a_1a_3}\rho_{a_2a_3}-\rho_{a_1a_2}}{\Delta},\ \Delta_{13} = \frac{\rho_{a_1a_2}\rho_{a_2a_3}-\rho_{a_1a_3}}{\Delta},\ \Delta_{23} = \frac{\rho_{a_1a_2}\rho_{a_1a_3}-\rho_{a_2a_3}}{\Delta},\\
\Lambda &= \Delta(\pi)^{-\frac{3}{2}}\left[(1-\rho^2_{a_1a_2})(1-\rho^2_{a_1a_3})(1-\rho^2_{a_2a_3})\right]^{-\frac{1}{2}},
\end{aligned}$$

and

$$\begin{aligned}
G_{1it} &= A_{1it} + a_m\sigma_{a_1}\sqrt{2(1-\rho^2_{a_1a_2})},\ G_{2it} = A_{2it} + a_p\sigma_{a_2}\sqrt{2(1-\rho^2_{a_1a_2})},\\
H_{1it} &= B_{1it} + \frac{a_m\sigma_{a_1}\sqrt{2}}{\sqrt{\Delta_{11}}},\ H_{2it} = B_{2it} + \frac{a_p\sigma_{a_2}\sqrt{2}}{\sqrt{\Delta_{22}}},\ H_{3it} = B_{3it} + \frac{a_q\sigma_{a_3}\sqrt{2}}{\sqrt{\Delta_{33}}}.
\end{aligned}$$

The product over i of the likelihood functions in equations (22) and (23) can be maximized using standard numerical procedures to obtain the maximum likelihood estimator of the VAR models. The covariance matrix of the estimator is obtained using standard hessian or outer product gradient methods.

As we use an unbalanced panel data, we need at least three observations over time for some enterprises, two of which need to be consecutive, to be able to identify the autoregressive parameters in eqs. (1)-(2) and (3)-(5), and those of the individual effects in eqs. (7)-(8) and (9)-(11). Adding enterprises for which only two consecutive observations are available, where the lagged variables and the initial conditions have the same value, increases the number of observations without harming the identification of the above parameters as long as we have some enterprises with at least three observations. Conditioning the likelihood on different initial conditions for all enterprises is acceptable if we assume to be in a steady state.[16] Whenever we include data for an enterprise for which no observations are available in the first wave, we condition the likelihood on the first observations available of the dependent variables.

[16]We did not reject the null hypothesis of equal coefficients for different initial conditions.

5 Results

Table 3 presents the estimates of the single input KPF when R&D is the only innovation input, and Table 4 presents the estimates with other innovation expenditures as the only innovation input. Table 5 presents the estimates of the double input KPF where both R&D and other innovation expenditures enter the knowledge production function. One of the main results of the three tables is that individual effects are relevant to innovation inputs and innovation output, regardless of the specification of the KPF and regardless of the sector. In all specifications and for both sectors, the model assuming the absence of individual effects is rejected at 1% significance level. For instance, the VAR model assuming the absence of individual effects using the specification of Table 3 has likelihood values of -4893.982 for the high-tech sector and -6810.728 for the low-tech sector. We now discuss other estimation results of the KPF regarding persistence, time lag and feedback effect, and simultaneity.

Persistence

As shown in all three tables, there is strong evidence of persistence in the intensity of innovation input, regardless of how it is measured, and in that of innovation output. Furthermore, we find strong evidence of persistence regardless of whether a single innovation input or two innovation inputs enter the knowledge production function. The evidence of persistence in the intensity of innovation input was to be expected as the analysis is conditional on enterprises being innovative in two successive waves of the CIS. Hence, there is an upward bias towards persistence in the conditional analysis that is to be removed by considering an unconditional analysis in which both innovative and non-innovative enterprises are to be included. The resulting analysis would be a panel data VAR model with selection and is to be included in a later version of the study. In both specifications of the KPF, the effect of the sum of initial and lagged R&D intensity on current R&D intensity is significantly larger (at 5% level) in the high-tech sector than in the low-tech sector (see Tables 3 and 5). We do not find such an evidence for other innovation input intensity (see Tables 4 and 5) and the share of innovative sales (see Tables 3-5).

Lag and feedback effect

The results suggest that the lag effect of innovation input on innovation output differs according to the type of innovation input (R&D versus non-R&D expenditures) and the type of the sector. More specifically, lagged R&D intensity has a positive and highly significant effect on current share of innovative sales in the high-tech sector, and plays no role in current share of innovative sales in the low-tech sector. As for lagged other innovation input intensity, it is irrelevant to current share

Table 3: Knowledge production function estimates with R&D as innovation input‡

Variable	Coefficient	(Std. Err.)	Coefficient	(Std. Err.)
	High-tech		Low-tech	
Current R&D intensity (in log)				
Past R&D intensity	0.122†	(0.064)	0.174**	(0.056)
Past share of innovative sales	0.188**	(0.048)	0.085†	(0.049)
Initial R&D intensity	0.326**	(0.053)	0.130**	(0.046)
Lagged size (in log)	-0.020	(0.068)	-0.018	(0.080)
Technology push (lagged)	0.567**	(0.171)	0.167	(0.186)
Demand pull (lagged)	0.034	(0.161)	0.056	(0.164)
Subsidy (lagged)	0.684**	(0.183)	0.691**	(0.168)
Intercept	-2.633**	(0.478)	-4.564**	(0.559)
Current share of innovative sales (in logit)				
Past R&D intensity	0.064**	(0.024)	0.006	(0.018)
Past share of innovative sales	0.211**	(0.056)	0.120*	(0.051)
Initial share of innovative sales	0.127*	(0.050)	0.179**	(0.044)
Lagged size (in log)	0.148**	(0.047)	-0.020	(0.046)
Technology push (lagged)	0.043	(0.121)	0.085	(0.108)
Demand pull (lagged)	-0.005	(0.114)	0.301**	(0.095)
Subsidy (lagged)	0.469**	(0.129)	0.206*	(0.097)
Intercept	-1.810**	(0.332)	-1.466**	(0.317)
Extra parameters				
σ_{a_1}	1.085**	(0.248)	1.234**	(0.333)
σ_{a_2}	0.716**	(0.065)	0.665**	(0.058)
$\rho_{a_1 a_2}$	-0.211	(0.253)	0.563**	(0.128)
σ_{ϵ_1}	2.018**	(0.408)	2.446**	(0.619)
σ_{ϵ_2}	1.455**	(0.128)	1.436**	(0.121)
$\rho_{\epsilon_1 \epsilon_2}$	0.281**	(0.051)	0.132**	(0.042)
Number of observations	1029		1345	
Log-likelihood	-4225.237		-5738.032	

‡Note: time and industry dummies are included in each equation.

Significance levels : †: 10% ∗: 5% ∗∗: 1%

of innovative sales regardless of the type of the sector. These results on the lag effect between innovation input and innovation output are similar whether a single innovation input or two innovation inputs enter the knowledge production function. In Table 5, we further report the lag effect estimate of each innovation input on the other. There seems to be a negative and significant lag effect of R&D intensity on current other innovation input intensity in both sectors. The reverse lag effect is also negative but hardly significant in the high-tech sector and insignificant in the low-tech sector. The results on the lag effect between innovation inputs suggests that enterprises tend to spend on the same type of innovation activities, hence tend not to increase the range of innovation expenditures over time. In other words, R&D and other innovation expenditures seem to be intertemporal substitutes.

The feedback effect of innovation output on innovation input also differs according to the type of innovation input and the type of the sector. More specifically, lagged share of innovative sales has a positive and significant effect on current R&D intensity, which is larger in terms of significance

Table 4: Knowledge production function estimates with other innovation expenditures as innovation input‡

Variable	Coefficient	(Std. Err.)	Coefficient	(Std. Err.)
	High-tech		Low-tech	
Current other innovation input intensity (in log)				
Past innovation input intensity	0.240**	(0.041)	0.264**	(0.041)
Past share of innovative sales	-0.047	(0.072)	-0.007	(0.063)
Initial innovation input intensity	0.032	(0.041)	-0.013	(0.039)
Lagged size (in log)	0.100	(0.100)	-0.290**	(0.103)
Technology push (lagged)	0.045	(0.266)	0.135	(0.246)
Demand pull (lagged)	0.167	(0.261)	0.317	(0.220)
Subsidy (lagged)	-0.912**	(0.281)	-0.884**	(0.217)
Intercept	-7.198**	(0.677)	-4.218**	(0.644)
Current share of innovative sales (in logit)				
Past innovation input intensity	-0.002	(0.015)	-0.005	(0.013)
Past share of innovative sales	0.205**	(0.056)	0.115*	(0.052)
Initial share of innovative sales	0.144**	(0.050)	0.185**	(0.045)
Lagged size (in log)	0.159**	(0.048)	-0.020	(0.046)
Technology push (lagged)	0.047	(0.120)	0.103	(0.107)
Demand pull (lagged)	0.020	(0.115)	0.293**	(0.095)
Subsidy (lagged)	0.554**	(0.124)	0.228*	(0.095)
Intercept	-2.237**	(0.309)	-1.543**	(0.286)
Extra parameters				
σ_{a_1}	0.354**	(0.047)	0.781**	(0.202)
σ_{a_2}	0.756**	(0.070)	0.682**	(0.059)
$\rho_{a_1 a_2}$	0.535*	(0.229)	0.576**	(0.148)
σ_{ϵ_1}	3.669**	(1.181)	3.550**	(1.213)
σ_{ϵ_2}	1.441**	(0.125)	1.429**	(0.119)
$\rho_{\epsilon_1 \epsilon_2}$	0.096*	(0.042)	-0.001	(0.038)
Number of observations	1029		1345	
Log-likelihood	-4745.518		-6157.549	

‡Note: time and industry dummies are included in each equation.
Significance levels : †: 10% *: 5% **: 1%

in the high-tech sector. There is no evidence of a feedback effect of the share of innovative sales on other innovation input intensity, regardless of the type of the sector.

Simultaneity

In all three tables, cross-equation "total" correlations between the disturbance terms are positively and significantly estimated.[17] One notable exception is the cross-equation correlation between the disturbances of R&D intensity and other innovation intensity (Table 5) that is insignificant in the high-tech sector, and negatively and significantly estimated in the low-tech sector. The results suggest strong evidence of simultaneity between innovation input and innovation output, regardless of the measure of innovation input and regardless of the specification of the KPF.

[17]The main expression of the "total" correlation between the disturbance terms μ_{jit} and μ_{kit} of equations j and k, denoted by $\rho_{u_j u_k}$ with $j \neq k$, is calculated as

$$\rho_{u_j u_k} = \frac{\rho_{a_j a_k}\sigma_{a_j}\sigma_{a_k} + \rho_{\epsilon_j \epsilon_k}\sigma_{\epsilon_j}\sigma_{\epsilon_k}}{\sqrt{(\sigma^2_{a_j} + \sigma^2_{\epsilon_j})(\sigma^2_{a_k} + \sigma^2_{\epsilon_k})}}.$$

Table 5: Knowledge production function estimates with R&D and other innovation expenditures as innovation inputs[‡]

Variable	Coefficient	(Std. Err.)	Coefficient	(Std. Err.)
	High-tech		Low-tech	
Current R&D intensity (in log)				
Past R&D intensity	0.123*	(0.063)	0.178**	(0.054)
Past innovation input intensity	-0.039†	(0.022)	-0.020	(0.023)
Past share of innovative sales	0.202**	(0.048)	0.092†	(0.050)
Initial R&D intensity	0.317**	(0.053)	0.126**	(0.045)
Lagged size (in log)	-0.023	(0.067)	-0.031	(0.080)
Technology push (lagged)	0.575**	(0.170)	0.179	(0.186)
Demand pull (lagged)	0.051	(0.161)	0.052	(0.164)
Subsidy (lagged)	0.693**	(0.183)	0.695**	(0.168)
Intercept	-2.951**	(0.507)	-4.620**	(0.571)
Current other innovation input intensity (in log)				
Past R&D intensity	-0.136*	(0.054)	-0.182**	(0.041)
Past innovation input intensity	0.231**	(0.041)	0.240**	(0.040)
Past share of innovative sales	-0.015	(0.073)	0.042	(0.064)
Initial innovation input intensity	0.028	(0.041)	-0.001	(0.038)
Lagged size (in log)	0.116	(0.100)	-0.239*	(0.103)
Technology push (lagged)	0.147	(0.269)	0.240	(0.245)
Demand pull (lagged)	0.233	(0.262)	0.350	(0.219)
Subsidy (lagged)	-0.670*	(0.294)	-0.683**	(0.221)
Intercept	-8.199**	(0.779)	-5.755**	(0.729)
Current share of innovative sales (in logit)				
Past R&D intensity	0.064**	(0.025)	0.008	(0.018)
Past innovation input intensity	0.003	(0.015)	-0.002	(0.013)
Past share of innovative sales	0.215**	(0.055)	0.116*	(0.050)
Initial share of innovative sales	0.123*	(0.049)	0.182**	(0.043)
Lagged size (in log)	0.149**	(0.047)	-0.024	(0.046)
Technology push (lagged)	0.043	(0.121)	0.090	(0.108)
Demand pull (lagged)	-0.004	(0.115)	0.296**	(0.095)
Subsidy (lagged)	0.468**	(0.130)	0.208*	(0.097)
Intercept	-1.792**	(0.353)	-1.447**	(0.325)
Extra parameters				
σ_{a_1}	1.060**	(0.233)	1.209**	(0.310)
σ_{a_2}	0.377**	(0.054)	0.762**	(0.175)
σ_{a_3}	0.706**	(0.063)	0.665**	(0.055)
$\rho_{a_1 a_2}$	-0.091	(0.806	0.434†	(0.252)
$\rho_{a_1 a_3}$	-0.196	(0.246)	0.530**	(0.139)
$\rho_{a_2 a_3}$	0.541*	(0.217)	0.542**	(0.160)
σ_{ϵ_1}	2.025**	(0.400)	2.456**	(0.600)
σ_{ϵ_2}	3.654**	(1.182)	3.530**	(1.130)
σ_{ϵ_3}	1.459**	(0.127)	1.435**	(0.115)
$\rho_{\epsilon_1 \epsilon_2}$	-0.048	(0.048)	-0.192**	(0.037)
$\rho_{\epsilon_1 \epsilon_3}$	0.276**	(0.050)	0.139**	(0.042)
$\rho_{\epsilon_2 \epsilon_3}$	0.101*	(0.041)	0.009	(0.037)
Number of observations	1029		1345	
Log-likelihood	-7012.958		-9353.521	

[‡]Note: time and industry dummies are included in each equation.

Significance levels : †: 10% ∗: 5% ∗∗: 1%

Finally, we find evidence of contemporaneous substitutability between R&D intensity and other innovation input intensity.

Strictly exogenous regressors

Regardless of the specification of the KPF, the following results show up. First, technology push has a positive and significant effect on the intensity of R&D in the high-tech sector. Secondly, demand pull has a positive and significant effect on the share of innovative sales in the low-tech sector. Thirdly, size enters negatively and significantly the equation of other innovation intensity, and positively and significantly the equation of the share of innovative sales. Finally, *ceteris paribus* subsidized enterprises have larger R&D intensity, smaller other innovation input intensity and larger share of innovative sales than non-subsidized counterparts.

6 Conclusion

We have studied the relationship between innovation input and innovation output using a knowledge production function along the lines of Pakes and Griliches (1980a,b). Several features of the KPF were accounted for, namely the persistence of innovation input and innovation output, the time lag and simultaneity between innovation input and innovation output, the feedback effect of innovation output on innovation input, and the importance of unobserved heterogeneity. We have considered two specifications of the KPF, namely with one and two innovation inputs. In the single input KPF, R&D expenditures and other non-R&D innovation expenditures interchange as innovation input, while they enter together the double input KPF. We have thus estimated by maximum likelihood two- and three-component VAR models with panel data, and have found the following results. First, individual effects play an important role in the KPF as the specifications assuming the absence of individual effects are all rejected at 1% level of significance. Secondly, there is evidence of strong persistence in the intensity of innovation input and in that of innovation output. Thirdly, the time lag between innovation input and innovation output differs according to the type of innovation input and the type of the sector. In other words, there is only a lag effect between R&D intensity and the share of innovative sales, and the lag is effective only in the high-tech sector. Fourthly, a similar result shows up for the feedback effect of innovation output on innovation input, i.e. lagged share of innovative sales has a positive and significant effect only on current R&D intensity. The effect is highly significant in the high-tech sector, while hardly significant in the low-tech sector. Finally, there is strong evidence of simultaneity between innovation input and innovation output in both sectors, regardless of how innovation input is measured, and intertemporal substitutability between both types of innovation input.

The main caveat of the study is that it is conditional on enterprises being innovative in two consecutive waves of the CIS, hence an upward bias towards the persistence in the intensity of innovation input. The next step is to extend the study including enterprises that are in two consecutive waves of the CIS, but not necessarily innovative in both waves. Hence, the panel data VAR model is to be extended so as to include a selection equation to explain the probability to be innovative.

References

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

David, P. A., B. H. Hall and A. A. Toole, "Is public R&D a Complement or Substitute for Private R&D? A Review of the Econometric Evidence," *Research Policy* 29 (2000), 497–529.

Hall, B. H., Z. Griliches and J. A. Hausman, "Patents and R and D: Is There a Lag?," *International Economic Review* 27 (1986), 265–283.

Hausman, J., B. H. Hall and Z. Griliches, "Econometric Models for Count Data with an Application to the Patents-R&D Relationship," *Econometrica* 52 (1984), 909–938.

Holtz-Eakin, D., W. Newey and H. S. Rosen, "Estimating Vector Autoregressions with Panel Data," *Econometrica* 56 (1988), 1371–1395.

MaCurdy, T. E., "Multiple Time Series Model Applied to Panel Data," NBER Working Paper No. 646, 1981.

———, "The Use of Time Series Processes to Model the Error Structure of Earnings in a Longitudinal Data Analysis," *Journal of Econometrics* 18 (1982), 83–114.

OECD, *Science, Technology and Industry Scoreboard. Benchmarking Knowledge-Based Economies* (Paris: OECD, 1999).

Pakes, A. and Z. Griliches, "Patents and R&D at the Firm Level: A First Look," NBER Working Paper No. 561, 1980a.

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

Raymond, W., *The Dynamics of Innovation and Firm Performance: An Econometric Panel Data Analysis*, Ph.D. thesis, Maastricht University (May 2007).

Wooldridge, J. M., "Simple Solutions to the Initial Conditions Problem in Dynamic Nonlinear Panel Data Models with Unobserved Heterogeneity," *Journal of Applied Econometrics* 20 (2005), 39–54.