Persistence of Innovation in Dutch Manufacturing: Is it Spurious?*

Wladimir Raymond[†], Pierre Mohnen[‡], Franz Palm[§] and Sybrand Schim van der Loeff[¶]

October 23, 2008

Abstract

This paper studies persistence of innovation in Dutch manufacturing using an unbalanced panel of firm data from four waves of the Community Innovation Survey between 1994 and 2002. We estimate by maximum likelihood a dynamic type 2 tobit model accounting for individual effects and handling the initial conditions problem. We find true persistence in the probability of innovating in the high-tech category of industries and spurious persistence in the low-tech category. Furthermore past innovation output intensity affects, albeit to a small extent, current innovation output intensity in the high-tech category, while no such evidence is found in the low-tech category.

Keywords: Dynamic type 2 tobit, Innovation, Panel data, Persistence.

JEL classification: C33, C34, O31

***Acknowledgements:** The empirical part of this study has been carried out at the Centre for Research of Economic Microdata at Statistics Netherlands. The authors wish to thank Statistics Netherlands, and in particular Bert Diederen, for helping them in accessing and using the Micronoom data set. The views expressed in this paper are solely those of the authors. The authors also wish to thank François Laisney and an anonymous referee for helpful comments. 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.

[†] University of Maastricht, W.Raymond@KE.unimaas.nl

[‡] University of Maastricht, MERIT and CIRANO, MERIT, University of Maastricht, P.O. Box 616 6200 MD Maastricht, The Netherlands; Tel: +31 43 388 3869; Fax: + 31 43 388 4905; P.Mohnen@MERIT.unimaas.nl

[§] University of Maastricht and CESifo fellow, F.Palm@KE.unimaas.nl

[¶] University of Maastricht, S.Loeff@KE.unimaas.nl

1 Introduction

This paper examines the persistence of innovation in Dutch manufacturing using an unbalanced panel of firm data from four waves of the Community Innovation Survey (henceforth CIS) pertaining to the periods 1994-1996, 1996-1998, 1998-2000 and 2000-2002. The main issue this paper addresses is whether success breeds success in innovation. We examine two aspects of persistence. First, does success in past innovation increase the probability of success in current innovation? Secondly, does past innovation output intensity, as measured by the share in total sales accounted for by sales of new or improved products (innovative sales),[1] positively affect current innovation output intensity?

Persistence of innovation plays an important role in endogenous growth and industrial dynamics. It can explain ongoing growth even in the absence of knowledge externalities. Furthermore, if innovation output intensity is closely linked to economic performance, as shown in the studies using the Crépon-Duguet-Mairesse type of model (Crépon et al., 1998), persistence in innovation output intensity may explain the persistence of firm economic performance such as productivity or profits (see Cefis and Ciccarelli, 2005).

Several theoretical explanations to the persistence of innovation have been put forward in the literature. One strand of literature following Schumpeter points to a relationship between market power and innovation. Monopolists have more to lose by not innovating than potential new entrants (see Gilbert and Newbery, 1982). Hence, incumbents tend to innovate persistently. A second explanation pertains to the financial constraints that a firm may face in funding its innovation activities. Indeed, because of information asymmetry between the innovator and the lender, the firm is unwilling to disclose valuable technological information that may be acquired by competitors (Bhattacharya and Ritter, 1983). Therefore, seeking external sources of funds may be more expensive than relying on retained earnings. As a result, profits that are generated by past successful innovations condition the financing of current innovation activities. A third explanation stems from the notion of technological trajectories defined in the evolutionary theory (see e.g. Nelson and Winter, 1982). Along a technological trajectory radical innovations are followed by a succession of incremental innovations. Consumers are inclined to buy new

[1] This study considers products that are new to the firm, but not necessarily new to the market.

generations of products increasing the demand for innovation. Furthermore, in a process similar to Arrow's learning-by-doing, firms learn by innovating and develop organizational competencies along the technological trajectory (Dosi and Marengo, 1994).

This study contributes to the empirical literature on innovation in a number of ways. First, it analyzes persistence using innovation output indicators from the CIS survey other than patents. Secondly, this is the first time that four waves of the CIS are used to investigate jointly the two aspects of persistence of innovation mentioned earlier, i.e. in qualitative as well as in quantitative terms. Third, we estimate a dynamic type 2 tobit model, according to Amemiya's (1984) terminology, using an unbalanced panel of firm data accounting for unobserved heterogeneity through individual effects. The incidence and the intensity of innovation are estimated jointly allowing for a correlation between the processes governing the introduction of new or significantly improved products and/or processes, and the generation of innovative sales. Both equations of the model follow a dynamic specification. We use an estimation technique suggested by Wooldridge (2005) to handle the initial conditions problem, and generalized in Raymond et al. (2007) to models with sample selection, and find true persistence in the incidence of innovation in the category of industries referred to as high-tech and spurious persistence in the category of industries referred to as low-tech. Furthermore past innovation output intensity affects, albeit to a small extent, current innovation output intensity in the high-tech category, while no such evidence is found in the low-tech category.

Section 2 summarizes the findings of the empirical literature on the persistence in the occurrence and in the intensity of innovation. We describe the data in Section 3, present the model in Section 4 and its estimation in Section 5. We present and discuss the estimation results in Section 6, and conclude in Section 7.

2 Literature

We briefly describe the empirical literature on the persistence in the occurrence and in the intensity of innovation. Some studies test the Schumpeter Mark I and II hypotheses, in other words whether innovation activities are subject to "creative destruction" or "creative accumulation" (Cefis and Orsenigo, 2001; Cefis, 2003). Other authors test whether innovation activities are subject to some kind of inertia, denoted by "dynamic economies of scales" or "success breeds success" (Crépon and Duguet, 1997; Geroski et al., 1997). In the evolutionary literature on industry dynamics, industry and country differences in the persistence of innovation are investigated (Malerba and Orsenigo, 1999).

The literature identifies two types of studies according to whether patent or other data are used. In the first type of studies innovation is measured by the number of patents granted or applied for at the European Patent Office (henceforth EPO) or the United States Patents and Trademarks Office (henceforth US PTO). In the second type of studies, innovation is measured on the input side by R&D or other innovation activities, or on the output side by the introduction of new products and/or processes on the basis of data from R&D or innovation surveys. Table 1 shows that all the studies on the persistence of innovation that use patent data, with the exception of Crépon and Duguet (1997), conclude that there is no clear-cut evidence of strong persistence in innovation activities, regardless of the methodology. In fact, those studies share a common drawback, namely the type of data used to analyze persistence. Indeed, in order for a firm to appear in a patent data set, it has to be the first to apply for a patent. Hence, when analyzing the persistence of innovation using patent data, one is unwittingly analyzing the persistence in "winning the patent race", which is unlikely to be strong.

Studies using major innovations yield results that are similar to those obtained with patent data (e.g. Geroski et al., 1997). Indeed, a major innovation is one that meets a path breaking success, which is unlikely to persist over a long period of time. On the contrary, R&D and innovation data allow persistence to be analyzed at the firm level without mentioning the patenting or market leadership status of the firm. In this case, regardless of the methodology, persistence in innovation activities is found to be high, whether input measures (Máñez Castillejo et al., 2004; Peters, 2005) or output measures (Flaig and Stadler, 1994; Duguet and Monjon, 2002) of innovation are used.

Most studies examine only the persistence in the decision to innovate using a dynamic probit or a duration model. Van Leeuwen's (2002) study is the only one that analyzes the dynamics of innovation input intensity and links it to that of innovation output intensity using two waves of the Dutch CIS data. However, his analysis does not account for individual effects, and does not model the dynamics in the decision to innovate.

Our study attempts to give a first insight into the link between the persistence of innovation and the dynamics of firms' innovation output intensity in Dutch manufacturing using four waves of the CIS.

3 Data

The data are collected by the *Centraal Bureau voor de Statistiek* (CBS) and stem from four waves of the Dutch CIS, 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 15.1-37.2) are included in the analysis.[2] The population of interest consists of enterprises with at least ten employees and positive sales at the end of the 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 wave of the CIS and PS resulting in about 3000 enterprises in each wave of the merged data of our sample.

In order to carry out our analysis, we consider enterprises that take part in at least two consecutive innovation and production surveys 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. The unbalanced panel is more representative of the population of interest than the balanced panel, contains more observations, and

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

controls in part for a survivorship bias that arises when using the balanced panel. We perform the analysis using both panels and contrast the results obtained to assess the magnitude of part of the survivorship bias. Since survival is related to size (Doms et al., 1995; Agarwal and Audretsch, 2001), the survivorship bias is not entirely accounted for because of the sampling of enterprises with less than 50 employees.

3.1 Dependent variables

As we examine two aspects of persistence, in the occurrence and in the intensity of innovation, we distinguish two dependent variables. The first dependent variable is binary indicating whether an enterprise is a technological product and/or process (TPP) innovator. In the innovation survey, an enterprise is asked (1) whether it has implemented at least one new or improved product, (2) whether it has implemented at least one new or improved process during the period under review. A TPP innovator is an enterprise that has responded positively to either (1) or (2) or both.

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.[3]

3.2 Explanatory variables

We explain the probability of being a current TPP innovator by the fact of having been a TPP innovator in the previous wave, lagged size, market share, being part of a group, industry dummies and time dummies. Measurements of these variables are available for both TPP and non-TPP innovators. Besides lagged size, being part of a group, industry dummies and time dummies, we explain current innovation output intensity by past innovation output intensity, three lagged R&D variables and lagged indicators for demand pull, proximity to science, cooperation and subsidy. All the additional explanatory variables of innovation output intensity stem from the CIS surveys. It is to be noted that a one period lag actually corresponds to two years.

[3]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.

According to the Schumpeterian tradition (Schumpeter, 1942), size and domestic market share influence positively the probability of being a TPP innovator. It is often argued that larger firms have better access to finance. They are more likely to engage in risky projects, and to benefit from economies of scale. Size is measured by the number of employees, and domestic market share is defined as the ratio of the sales of an enterprise over the total sales of the 3-digit industry it belongs to.[4] The number of employees and sales stem from the PS and are measured for the last year of the period under review. Size and domestic market share are log-transformed in the estimation.

Firms that are part of a group are expected to be more innovative because they benefit from knowledge spillovers, internal access to finance, and synergies in marketing (Veugelers and Cassiman, 2004). A binary variable indicating whether an enterprise belongs to a group during the period under review is directly reported in the CIS.

We include three R&D variables as explanatory variables on the grounds that experience and knowledge accumulated from past R&D have a positive influence on innovation output as in the knowledge production function literature (Hall et al., 1986). The first, R&D intensity, is the ratio of total (intramural and extramural) R&D expenditures over total sales. This variable is measured at the end of the period under review. Its logarithmic transformation is used in the estimation when the variable takes on a positive value. Secondly, a dummy variable for non-R&D performer is included in the analysis to "compensate" for the fact that the log transformation of the original R&D intensity is set to zero for observations with original R&D intensity equal to zero. The third R&D variable is continuous R&D that takes on the value one if the enterprise reports that it performed intramural R&D continuously during the period under review, and zero otherwise.

Following Schmookler (1966) innovation is driven by demand. Most empirical studies find a positive impact of demand pull on the share of innovative sales, regardless of the proxy used for demand pull. 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".

[4] 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.

Other authors credit technology as the major source of innovation (Rosenberg, 1974). Proximity to science is proxied by a dummy variable constructed from the indicator 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.

Enterprises that undertake innovative activities in cooperation are expected to benefit from knowledge spillovers, hence to perform better technologically (d'Aspremont and Jacquemin, 1988). A dummy variable for cooperation is introduced that takes on the value one if the enterprise reports that it undertook its innovative activities in cooperation, and zero otherwise.

We expect enterprises that receive subsidies for innovation to be more innovative, although evidence on this score is mixed (David et al., 2000). 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.

In addition to unobserved heterogeneity captured by individual effects in the two equations of the model, we also account for industry and time effects by introducing appropriate dummies in each equation.

3.3 Descriptive statistics and transition probabilities

Table 2 shows descriptive statistics of the dependent and explanatory variables for both the unbalanced and the balanced panel. For instance, 65% and 68% of the enterprises in the unbalanced and balanced panel respectively are TPP innovators. Furthermore, the average share of innovative sales of TPP innovators is 28% in the unbalanced panel and 29% in the balanced panel. There is a slightly higher proportion of TPP innovators and innovative sales in the balanced panel, suggesting a positive link, between innovation and survival. But, it may also be due to a selection bias. Indeed, even our unbalanced panel is biased towards large firms because of the sampling procedure and towards firms that have survived at least 5 years.[5] The enterprises of the balanced panel are on average larger and

[5]The literature emphasizes the fact that few new firms survive the critical start-up period (Mata and Portugal, 1994).

more homogenous in terms of size than those of the unbalanced panel. Apart from size the enterprise characteristics are on average similar across panels.

Table 3 reports transition probabilities from period t-1 to period t for both the unbalanced and the balanced panel. The upper part of the table shows transition probabilities for the innovation status. For instance, in the unbalanced panel, 59% of non-TPP innovators and 81% of TPP innovators in CIS 2 remain in their initial state in CIS 2.5. The corresponding figures are 56% and 85% in the balanced panel. The lower part of the table shows transition probabilities between the states of being below and above the average share of innovative sales in two successive waves. For instance, 49% and 81% of the innovators with respectively below and above average share of innovative sales in CIS 2 remain in their initial state in CIS 2.5. The corresponding figures are 53% and 85% in the balanced panel. The general pattern of Table 3 is that TPP innovation status and innovation output intensity are fairly persistent, which may be due to true or spurious state dependence. In order to distinguish the former from the latter, we consider a model of innovative behavior in a dynamic panel data framework that accounts for unobserved individual effects that are correlated with the initial conditions. The model is a dynamic panel data type 2 tobit which encompasses the cross-sectional type 2 tobit model studied by, for instance, Mairesse and Mohnen (2001).

4 Econometric model

The model explains the occurrence of TPP innovations in Dutch manufacturing enterprises and the extent of these innovations in terms of the share of innovative sales. Formally, it is written as

$$d_{it} = \mathbf{1}[\rho d_{i,t-1} + \delta' \mathbf{w}_{it} + \eta_i + \epsilon_{1it} > 0] \tag{1}$$

$$y_{it} = \begin{cases} \gamma y_{i,t-1} + \beta' \mathbf{x}_{it} + \alpha_i + \epsilon_{2it} & \text{if } d_{it} = 1, \\ 0 & \text{if } d_{it} = 0, \end{cases} \tag{2}$$

where $t = 1, ...T_i$, $i = 1, ...N$, and $\mathbf{1}[...]$ is an indicator function that takes on the value one if the expression between square brackets is true, and zero otherwise.

Equation (1) models the current decision of enterprise i to innovate as a latent function of its past innovation achievement ($d_{i,t-1}$), its observable characteristics ($\mathbf{w}_{it}$),[6] time-invariant unobserved individual effects (η_i) and other time-variant unobserved variables (ϵ_{1it}) uncorrelated with $\mathbf{w}_{it}$. The expression in square brackets represents the incentive to innovate. If the incentive is sufficiently high, enterprise i is a TPP innovator in which case d_{it} is observed to be 1. We include in $\mathbf{w}_{it}$ lagged size and domestic market share rather than their current counterparts so as to avoid explaining the probability of being a TPP innovator during a given survey period by explanatory variables measured at the end of that period. The scalar ρ and the vector δ' capture respectively the effects of past innovation achievement and firm characteristics on current innovation achievement and are to be estimated. A positive and statistically significant estimate of ρ identifies the presence of persistence in the occurrence of innovation which may occur for two reasons, because of state dependence or because of unobserved effects or left-out variables that are correlated over time (through serially-correlated errors or individual effects). Heckman (1981) refers to the first phenomenon as true state dependence and the second one as spurious state dependence. True state dependence states that past innovation achievement increases positively and significantly the probability of current innovation achievement (true persistence).[7] In order to distinguish it from spurious state dependence, unobserved effects that are correlated over time and the endogeneity of the initial conditions must be properly accounted for when estimating eq. (1).

Equation (2) models the current share of innovative sales (y_{it}) of innovator i ($d_{it} = 1$) as being determined by its past share of innovative sales ($y_{i,t-1}$), its characteristics ($\mathbf{x}_{it}$), time-invariant unobserved individual effects (α_i) and other time-variant unobserved variables (ϵ_{2it}) uncorrelated with $\mathbf{x}_{it}$. This share is zero if enterprise i is not an innovator, and the full set of regressors included in $\mathbf{x}_{it}$ are only available when enterprise i is an innovator. Besides lagged size and being part of a group, we include in $\mathbf{x}_{it}$ three lagged R&D variables, and lagged indicators for demand pull, proximity to science, cooperation and subsidy. We allow for a one period lag between innovation determinants and innovation output intensity. The scalar γ and the vector β' capture respectively the effects of past share of innovative sales and firm characteristics on current share of innovative sales and are to be

[6] $\mathbf{w}_{it}$ could also include market specific characteristics if they were observable.

[7] When the term persistence is used in this study without any further explanation, it is to be understood as true persistence.

estimated.

Equations (1) and (2) are jointly estimated allowing for a correlation between the processes governing the introduction of TPP innovations and the generation of innovative sales.

5 Maximum likelihood estimation

The econometric literature on dynamic panel data shows that the coefficient of the lagged dependent variable is overestimated when individual effects and the initial conditions are not properly accounted for. Estimation techniques that properly handle these problems in nonlinear dynamic panel data models are known in the literature (Wooldridge, 2005).

Semi-parametric fixed-effects approaches along the lines of Kyriazidou (2001) cannot be applied to our data, since they consist mostly of qualitative variables and the few continuous ones that we have exhibit too little within variation. Hence, most of the variables would be wiped out when taking first differences. For instance, we would not be able to identify the effects of industry dummies that are assumed to capture technological opportunities.

In order to cope with the characteristics of our data, we consider an error-components approach and make distributional assumptions on the individual effects. We "integrate out" the individual effects and use the Wooldridge (2005) approach of handling the initial conditions problem. The estimator is described as follows. We assume the individual effects to be correlated with the initial conditions and the regressors, i.e.

$$\eta_i = b_0^s + b_1^s d_{i0} + \mathbf{b}_2'^s \mathbf{w}_i + a_{1i}, \tag{3}$$

$$\alpha_i = b_0^r + b_1^r y_{i0} + \mathbf{b}_2'^r \mathbf{x}_i + a_{2i}, \tag{4}$$

where $\mathbf{w}_i' = (\mathbf{w}_{i1}', ..., \mathbf{w}_{iT_i}')$, $\mathbf{x}_i' = (\mathbf{x}_{i1}', ..., \mathbf{x}_{iT_i}')$, d_{i0} and y_{i0} pertain to the first available observation for each enterprise, and b_0^s, b_1^s, $\mathbf{b}_2'^s$, b_0^r, b_1^r and $\mathbf{b}_2'^r$ are to be estimated.[8] The scalars b_1^s and b_1^r capture the dependence of the individual effects on the initial conditions.

[8]The approach considered in equations (3) and (4) allows the individual effects to be correlated with the regressors. However, because of the lack of variation over time (within variation) in $\mathbf{w}_{it}$ and $\mathbf{x}_{it}$, we are not able to identify δ' from $\mathbf{b}_2'^s$ and β' from $\mathbf{b}_2'^r$, and therefore we assume the individual effects to be correlated only with the initial conditions.

The vectors $(a_{1i}, a_{2i})'$ and $(\epsilon_{1it}, \epsilon_{2it})'$ are assumed to be independently and identically (over time and across individuals) normally distributed with means zero and covariance matrices $\Omega_{a_1a_2} = \begin{pmatrix} \sigma^2_{a_1} & \rho_{a_1a_2}\sigma_{a_1}\sigma_{a_2} \\ \rho_{a_1a_2}\sigma_{a_1}\sigma_{a_2} & \sigma^2_{a_2} \end{pmatrix}$ and $\Omega_{u\epsilon} = \begin{pmatrix} 1 & \rho_{\epsilon_1\epsilon_2}\sigma_{\epsilon_2} \\ \rho_{\epsilon_1\epsilon_2}\sigma_{\epsilon_2} & \sigma^2_{\epsilon_2} \end{pmatrix}$ respectively, and independent of each other. The likelihood function of one individual, starting from $t = 1$ and conditional on the regressors and the initial conditions, is written as

$$L_i = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\prod_{t=1}^{T_i} L_{it}(d_{it}, y_{it}|d_{i0}, d_{i,t-1}, \mathbf{w}_i, y_{i0}, y_{i,t-1}, \mathbf{x}_i, a_{1i}, a_{2i})g(a_{1i}, a_{2i})da_{1i}da_{2i}, \quad (5)$$

where $\prod_{t=1}^{T_i} L_{it}(d_{it}, y_{it}|d_{i0}, d_{i,t-1}, \mathbf{w}_i, y_{i0}, y_{i,t-1}, \mathbf{x}_i, a_{1i}, a_{2i})$ is the likelihood function of individual i conditional on the individual effects, and $g(a_{1i}, a_{2i})$ is the bivariate normal density function of $(a_{1i}, a_{2i})'$. Define

$$A_{it} = \rho d_{i,t-1} + \delta'\mathbf{w}_{it} + b_0^s + b_1^s d_{i0} + \mathbf{b}_2'^s\mathbf{w}_i, \quad (6)$$

$$B_{it} = \gamma y_{i,t-1} + \beta'\mathbf{x}_{it} + b_0^r + b_1^r y_{i0} + \mathbf{b}_2'^r\mathbf{x}_i, \quad (7)$$

the likelihood function of individual i conditional on the individual effects is written as

$$\prod_{t=1}^{T_i}\Phi\left[-(A_{it} + a_{1i})\right]^{(1-d_{it})}\left[\frac{1}{\sigma_{\epsilon_2}}\phi\left(\frac{y_{it}-B_{it}-a_{2i}}{\sigma_{\epsilon_2}}\right)\Phi\left(\frac{A_{it} + a_{1i} + \frac{\rho_{\epsilon_1\epsilon_2}}{\sigma_{\epsilon_2}}(y_{it}-B_{it}-a_{2i})}{\sqrt{1-\rho^2_{\epsilon_1\epsilon_2}}}\right)\right]^{d_{it}}. \quad (8)$$

The double integral in equation (5) can be approximated by "two-step" Gauss-Hermite quadrature which states that

$$\int_{-\infty}^{\infty} e^{-z^2}f(z)dz \simeq \sum_{m=1}^{M} w_m f(a_m), \quad (9)$$

where w_m and a_m are respectively the weights and abscissas of the Gauss-Hermite integration, 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.

Equation (5) is written as

$$L_i = \int_{-\infty}^{\infty} l(a_{2i}) \prod_{t=1}^{T_i} \left[\frac{1}{\sigma_{\epsilon_2}} \phi \left(\frac{y_{it} - B_{it} - a_{2i}}{\sigma_{\epsilon_2}} \right) \right]^{d_{it}} H(a_{2i}) da_{2i}, \tag{10}$$

where $H(a_{2i})$ is written as

$$\int_{-\infty}^{\infty} m(a_{1i}, a_{2i}) \prod_{t=1}^{T_i} \Phi \left[- \left(A_{it} + a_{1i} \right) \right]^{(1-d_{it})} \left[\Phi \left(\frac{A_{it} + a_{1i} + \frac{\rho_{\epsilon_1\epsilon_2}}{\sigma_{\epsilon_2}} (y_{it} - B_{it} - a_{2i})}{\sqrt{1-\rho^2_{\epsilon_1\epsilon_2}}} \right) \right]^{d_{it}} da_{1i}, \tag{11}$$

where $l(a_{2i})$ and $m(a_{1i}, a_{2i})$ are functions of the respective arguments. In the first step, we approximate equation (11) using eq. (9). In the second step, we replace the approximation into eq. (10) and apply again eq. (9). The final expression of the likelihood is written as

$$\begin{aligned} L_i \simeq & \frac{\sqrt{1 - \rho^2_{a_1a_2}}}{\pi} \sum_{p=1}^{P} w_p \left\{ \prod_{t=1}^{T_i} \left[\frac{1}{\sigma_{\epsilon_2}} \phi \left(\frac{y_{it} - B_{it} - a_p \sigma_{a_2} \sqrt{2(1 - \rho^2_{a_1a_2})}}{\sigma_{\epsilon_2}} \right) \right]^{d_{it}} \right. \\ & \times \sum_{m=1}^{M} w_m \left\{ \exp \left[2\rho_{a_1a_2} a_p a_m \right] \prod_{t=1}^{T_i} \Phi \left[- \left(A_{it} + a_m \sigma_{a_1} \sqrt{2(1 - \rho^2_{a_1a_2})} \right) \right]^{(1-d_{it})} \right. \\ & \left. \left. \times \Phi \left(\frac{A_{it} + a_m \sigma_{a_1} \sqrt{2(1 - \rho^2_{a_1a_2})} + \frac{\rho_{\epsilon_1\epsilon_2}}{\sigma_{\epsilon_2}} (y_{it} - B_{it} - a_p \sigma_{a_2} \sqrt{2(1 - \rho^2_{a_1a_2})})}{\sqrt{1-\rho^2_{\epsilon_1\epsilon_2}}} \right)^{d_{it}} \right\} \right\}, \end{aligned} \tag{12}$$

where w_m, w_p, a_m and a_p are respectively the weights and abscissas of the first- and second-stage Gauss-Hermite integration with M and P being the first- and second-stage total number of integration points.[9] The same number of integration points ($P = M$) is used in this study, although P need not be equal to M. Equations (1) and (2) are correlated through the individual effects ($\rho_{a_1a_2} \neq 0$) and the idiosyncratic errors ($\rho_{\epsilon_1\epsilon_2} \neq 0$), and the "total" correlation between the two equations is calculated as

$$\rho_{tot} = \frac{\rho_{a_1a_2}\sigma_{a_1}\sigma_{a_2} + \rho_{\epsilon_1\epsilon_2}\sigma_{\epsilon_2}}{\sqrt{(\sigma^2_{a_1} + 1)(\sigma^2_{a_2} + \sigma^2_{\epsilon_2})}}. \tag{13}$$

When we use the unbalanced panel data set, we need at least three observations over time for some of the firms, two of which need to be consecutive, to be able to identify the

[9] Details on the calculation of the double integral can be found in Raymond et al. (2007).

parameters of the lagged dependent variables in eqs. (1) and (2) and those of the individual effects in eqs. (3) and (4). Adding firms 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 firms with at least three observations. Conditioning the likelihood on different initial conditions for all firms is acceptable if we assume to be in a steady state.[10] Whenever we include data for a firm for which no observations are available in the first wave, in expression (5) we condition on the first observations available of d_{i0} and y_{i0}.

6 Results

As product-life cycle varies across industries, the persistence of innovation may be expected to be industry-specific (Malerba and Orsenigo, 1999; Cefis and Orsenigo, 2001). We regrouped the 3-digit industries into 4 categories following the OECD classification, namely high-technology (HT), medium-high technology (MHT), medium-low technology (MLT) and low-technology (LT) industries.[11] We estimated the model by interacting the lagged dependent variable of each equation ($d_{i,t-1}$, $y_{i,t-1}$) with four dummies for the industry categories and allowing for different industry category intercepts. Then a Wald test was performed on the equality of the coefficients of the lagged innovation variables, the persistence parameters, across industry categories.[12] With a $\chi^2_{(4)} = 5.340$ and a p-value=0.254 we could not reject the equality between HT and MHT, and between MLT and LT. Let us call the two resulting categories high-tech and low-tech.[13] In Table 4 the results of the consequent model are presented.

In order to show the importance of accounting for individual effects and handling the

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

[11] The OECD classification (OECD, 1999) distinguishes four groups of industries: high-technology industries (SBI 24.4, 30.0, 32.1-32.3, 33.1-33.5, and 35.3), medium-high-technology industries (SBI 24.1-24.3, 24.5-24.7, 29.1-29.7, 31.1-31.6, 34.1-34.3, 35.2, and 35.5), medium-low-technology industries (SBI 23.1-23.3, 25.1, 25.2, 26.1-26.8, 27.1-27.5, 28.1-28.7, and 35.1), and low-technology industries (15.1-15.9, 16.0, 17.1-17.7, 18.1-18.3, 19.1-19.3, 36.1-36.6, 37.1, and 37.2).

[12] The test was performed on the model accounting for individual effects correlated with the initial conditions using the unbalanced panel.

[13] We have also experimented with 2-digit industry dummies, and concluded on the basis of a Wald test that the industries could also be regrouped into a high-tech and a low-tech category as far as the persistence parameters are concerned.

initial conditions problem, we present estimation results for three variants of the dynamic type 2 tobit model using the unbalanced panel in the first three pairs of columns in Table 4. More specifically, we present the estimation results of the model without accounting for individual effects in the first pair of columns, and those of the same model with individual effects taken into account but the initial conditions assumed to be exogenous in the second pair of columns.[14] These results are to be contrasted with the estimates in the third pair of columns resulting from the estimation of the model with individual effects correlated with the initial conditions. We also report estimation results for the third variant of the model using the balanced panel in the last pair of columns. A difference in the estimates of the model using the balanced and unbalanced panel is a partial indication of the magnitude of the survivorship bias.

The estimation results on the persistence in the occurrence of innovation, i.e. the estimates of the parameters in equation (1), are presented in the upper part of Table 4 and discussed in subsection 6.1. The estimation results on the persistence in the intensity of innovation, i.e. the estimates of the parameters in equation (2), are presented in the middle part of Table 4 and discussed in subsection 6.2. The lower part of Table 4 shows the estimates of the coefficients of the initial conditions, the standard deviations of the individual effects and the cross-equation correlations.

6.1 Persistence in the occurrence of innovation

The estimation results of the dynamic type 2 tobit model in the absence of individual effects and in the presence of individual effects but exogenous initial conditions are very similar. The persistence parameter in each category of industries is positive and highly significant, and lagged size, lagged domestic market share and being part of a group affect positively and significantly the probability to innovate. As mentioned earlier, persistence of innovation may be spurious. The existence of true persistence can only be ascertained after accounting for individual effects and handling properly the initial conditions. Once this is done, the hypothesis that the persistence parameter is equal to zero can no longer be rejected in the low-tech category of industries, while the persistence parameter remains significant in the high-tech category. This result contrasts with that of Duguet and

[14]Exogenous initial conditions imply that they are uncorrelated with the individual effects.

Monjon (2002) who find strong persistence in achieving TPP innovations in all French manufacturing industries. However, they do not account for individual effects.

We find, however, a significant persistence in the occurrence of innovation in high-tech industries in sharp contrast to almost all the previous studies that use patent data.[15] The finding that persistence is only present in high-tech industries is consistent with the findings of Blundell et al. (1999), Aghion et al. (2005) and Acemoglu et al. (2006). They find that industries that are closer to the "technological frontier" are more competitive and that the competitive pressure pushes firms to innovate. Following this logic, firms in high-tech industries are in general closer to the "technological frontier" and therefore they are more likely to display persistence in innovation. The estimates are more precise in the unbalanced panel but not statistically different from those of the balanced panel at the 5% significance level.

6.2 Persistence in the intensity of innovation

The estimates of the parameters of equation (2) are again similar in the model without individual effects and in the model with individual effects but exogenous initial conditions. The persistence parameter of each category of industries is positive and highly significant. Furthermore, lagged size, lagged R&D intensity, and lagged subsidy affect positively the current share of innovative sales and, *ceteris paribus*, past non-R&D performers perform worse in terms of innovation output intensity than past R&D performers.

After accounting for individual effects that are correlated with the initial conditions, the persistence parameter remains significant at the 5% level, but small in magnitude, in the high-tech category, but is no longer significant in the low-tech category. This means that the past share of innovative sales affects the current share of innovative sales in the high-tech category. Again the estimates are more precise in the unbalanced panel but not statistically different from those of the balanced panel at the 5% significance level. The similarity of the estimates confirms our suspicion that using the unbalanced panel is not sufficient to correct for the survivorship bias because it is itself biased towards large firms

[15] Part of this persistence may be due to a one-year overlap between two consecutive innovation surveys. If we assume the probability to innovate to be uniformly distributed across the three years covered by each survey, there would be roughly a 10% chance that an enterprise would declare to be innovative in two consecutive surveys while only innovating in the overlapping year. This magnitude, however, is not sufficient to explain the entire persistence found in high-tech industries.

(more than 50 employees) and towards firms that survive at least five years.[16]

Both the model that assumes the absence of individual effects and the one that accounts for individual effects but assumes exogenous initial conditions are rejected at the 1% level of significance by a likelihood ratio test. Hence, the full model is the preferred one where equations (1) and (2) are jointly estimated allowing for a correlation between the processes governing the introduction of TPP innovations and the generation of innovative sales. The two equations are found to be correlated, mainly through the individual effects, and the cross-equation "total" correlation (eq. (13)) is calculated ex post to be 0.265 for the unbalanced panel and 0.230 for the balanced panel.

7 Conclusion

This study gives first insights into the joint persistence of innovation occurrence and innovation output intensity in Dutch manufacturing using four waves of the Community Innovation Survey. We estimate a dynamic type 2 tobit model and find true persistence of innovation in the category of industries referred to as high-tech but only spurious persistence in the other category of industries referred to as low-tech. Furthermore, past innovation output intensity affects, albeit to a small extent, current innovation output intensity in the high-tech category, while no such evidence is found in the low-tech category. According to our results, there is persistence in innovation output when innovation is measured by the appearance of new products and/or processes and the eventual share in total sales due to new products, at least in enterprises that belong to the high tech category. Previous studies with one exception have found no persistence in patenting. Our results also contrast with those of Duguet and Monjon (2002), who find evidence of strong persistence of innovation in all French manufacturing industries. Our results are in accordance with the conclusions of Blundell et al. (1999) and Aghion et al. (2005) that firms which belong to industries that are more competitive and closer to the "technological frontier" have more incentives to innovate, hence tend to innovate persistently. The individual effects and their correlation with the initial conditions are important to account for when estimating the introduction of TPP innovations and the generation of

[16] We thank an anonymous referee for pointing this out.

innovative sales. Both processes are shown to be positively and significantly correlated, mainly through the individual effects.

Our results confirm the inherent characteristics of the innovation process identified by economic theory. First, the process is dynamic and should be derived from an intertemporal maximization problem. Secondly, differences in innovation behavior cannot be solely attributed to observable differences across firms (e.g. high-tech versus low-tech). Unobserved heterogeneity, through individual effects, plays a crucial role in accounting for differences in innovation behavior and must be modeled. Finally, qualitative and quantitative measures of innovation (output) must be modeled jointly as they are closely related to one another.

The main caveat of this study is the data we use to implement the model. First, the panel is rather short ($T = 4$), which may explain in part the lack of true persistence in the low-tech category of industries. Secondly, there is one-year overlap between two consecutive waves of the Dutch CIS. Hence, to the extent that respondents answer this survey consistently, the overlap would tend to bias the results towards persistence in being a TPP innovator and in innovation output intensity. However, as no evidence of persistence is found in the preferred model for low-tech, it may be concluded that the effect of the overlapping year is not important.

References

Abramovitz, Milton, and Irene 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, Daron, Philippe Aghion, and Fabrizio Zilibotti (2006) "Distance to Frontier, Selection, and Economic Growth," *Journal of the European Economic Association* 4, 37–74.

Agarwal, Rajshree, and David B. Audretsch (2001) "Does Entry Size Matter? The Impact of the Life Cycle and Technology on Firm Survival," *The Journal of Industrial Economics* 49, 21–43.

Aghion, Philippe, Nick Bloom, Richard Blundell, Rachel Griffith, and Peter Howitt (2005) "Competition and Innovation: An Inverted-U Relashionship," *Quaterly Journal of Economics* 120, 701–728.

Amemiya, Takeshi (1984) "Tobit Models: A Survey," *Journal of Econometrics* 24, 3–62.

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

Blundell, Richard, Rachel Griffith, and John van Reenen (1999) "Market Share, Market Value and Innovation in a Panel of British Manufacturing Firms," *Review of Economic Studies* 66, 529–554.

Cefis, Elena (2003) "Is there Persistence in Innovative Activities?," *International Journal of Industrial Organization* 21, 489–515.

Cefis, Elena, and Luigi Orsenigo (2001) "The Persistence of Innovative Activities A Cross-Countries and Cross-Sectors Comparative Analysis," *Research Policy* 30, 1139–1158.

Cefis, Elena, and Matteo Ciccarelli (2005) "Profit Differentials and Innovation," *Economics of Innovation and New Technology* 14, 43–61.

Crépon, Bruno, and Emmanuel Duguet (1997) "Estimating the Innovation Function from Patent Numbers: GMM on Count Panel Data," *Journal of Applied Econometrics* 12, 243–263.

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

d'Aspremont, Claude, and Alexis Jacquemin (1988) "Cooperative and Noncooperative R&D in Duopoly with Spillovers," *American Economic Review* 78, 1133–1137.

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

Doms, Mark, Timothy Dunne, and Mark J. Roberts (1995) "The Role of Technology Use in the Survival and Growth of Manufacturing Plants," *International Journal of Industrial Organization* 13, 523–542.

Dosi, Giovanni, and Luigi Marengo (1994) "Some Elements of an Evolutionary Theory of Organizational Competencies," in Richard W. England (Ed.), *Evolutionary Concepts in Contemporary Economics* (Ann Arbor: Michigan University Press).

Duguet, Emmanuel, and Stéphanie Monjon (2002) "Creative Destruction and the Innovative Core: Is Innovation Persistent at the Firm Level?" UCL Discussion Paper 02-07.

Flaig, Gebhard, and Manfred Stadler (1994) "Success Breeds Success. The Dynamics of the Innovation Process," *Empirical Economics* 19, 55–68.

Geroski, Paul A., John van Reenen, and Chris F. Walters (1997) "How Persistently do Firms Innovate?," *Research Policy* 26, 33–48.

Gilbert, Richard J., and David M. G. Newbery (1982) "Preemptive Patenting and the Persistence of Monopoly," *American Economic Review* 72, 514–526.

Hall, Bronwyn H., Zvi Griliches, and Jerry A. Hausman (1986) "Patent and R and D: Is There a Lag?," *International Economic Review* 27, 265–283.

Heckman, James J. (1981) "Statistical Models for Discrete Panel Data," in Charles F. Manski, and Daniel McFadden (Eds.), *Structural Analysis of Discrete Data with Econometric Applications* (Cambridge, MA: MIT Press) pp. 114–178.

Kyriazidou, Ekaterini (2001) "Estimation of Dynamic Panel Data Sample Selection Models," *Reviews of Economic Studies* 68, 543–572.

Máñez Castillejo, Juan A., María E. Rochina Barrachina, Amparo Sanchis Llopis, and Juan A. Sanchis Llopis (2004) "A Dynamic Approach to the Decision to Invest in R&D: The Role of Sunk Costs." mimeo.

Mairesse, Jacques, and Pierre Mohnen (2001) "To Be or Not To Be Innovative: An Exercise in Measurement," *STI Review Special Issue on New Science and Technology Indicators* 27, 103–129.

Malerba, Franco, and Luigi Orsenigo (1999) "Technological Entry, Exit and Survival: an Empirical Analysis of Patent Data," *Research Policy* 28, 643–660.

Mata, Jose, and Pedro Portugal (1994) "Life Duration of New Firms," *The Journal of Industrial Economics* 42, 227–245.

Nelson, Richard R., and Sidney G. Winter (1982) *An Evolutionary Theory of Economic Change* (Cambridge, MA: Harvard University Belnap Press).

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

Peters, Bettina (2005) "Persistence of Innovation: Stylised Facts and Panel Data Evidence." ZEW Discussion Paper, No. 05-81.

Raymond, Wladimir, Pierre Mohnen, Franz Palm, and Sybrand Schim van der Loeff (2007) "The Behavior of the Maximum Likelihood Estimator of Dynamic Panel Data Sample Selection Models." CESifo Working Paper No. 1992.

Rosenberg, Nathan (1974) "Science, Invention and Economic Growth," *The Economic Journal* 84, 90–108.

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

Schumpeter, Joseph A. (1942) *Capitalism, Socialism and Democracy* (New York: Harper and Brothers).

van Leeuwen, George (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.

Veugelers, Reinhilde, and Bruno Cassiman (2004) "Foreign Subsidiaries as a Channel of International Technology Diffusion: Some Direct Firm Level Evidence from Belgium," *European Economic Review* 48, 455–476.

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

Table 1: Empirical studies on the persistence of innovation

Study	Country and (Time-period)	Innovation activities	Methodology	Measure of persistence	Result
			Patent data		
Crépon and Duguet (1997)	France (1984-1989)	patents applied for at the EPO	GMM on dynamic count panel data model	effects of lagged patents	high persistence
Geroski et al. (1997)	UK (1969-1988)	patents granted by the US PTO	duration dependence Weibull model	length of innovation spell	low persistence
Malerba and Orsenigo (1999)	France, Germany Italy, Japan, UK, US (1978-1991)	patents applied for at the EPO	descriptive analysis	duration of patenting after entry	low persistence
Cefis and Orsenigo (2001)	France, Germany Italy, Japan, UK, US (1978-1993)*	patents applied for at the EPO	TPM‡used in 1^{st} and 2^{nd} order Markov chains	probability of remaining in the same state of patenting	bimodality† and low persistence
Cefis (2003)	UK (1978-1991)	patents applied for at the EPO	TPM used in 1^{st} and 2^{nd} order Markov chains	probability of remaining in the same state of patenting	bimodality and low persistence
			R&D and innovation survey data		
Flaig and Stadler (1994)	West Germany (1979-1986)	achieve product and/or process innovations	Heckman ML estimation on RE** dynamic probit with panel data	lagged product and/or process innovations	high persistence
Geroski et al. (1997)	UK (1945-1982)	achieve at least one major innovation	duration dependence Weibull model	length of innovation spell	low persistence
Duguet and Monjon (2002)	France (1986-1996)‡‡	achieve product and/or process innovations (CIS)	ML estimation on dynamic probit with no individual effects	lagged product and/or process innovations	high persistence
Máñez Castillejo et al. (2004)	Spain (1990-2000)	engage in R&D activities	SML on SICM dynamic probit with panel data††	lagged R&D activities	high persistence
Peters (2005)	Germany (1994-2002)	engage in innovation activities (CIS)	Wooldridge ML estimation on RE dynamic probit with panel data	lagged innovation activities	high persistence

*The period is 1978-1991 for the UK. ‡TPM means transition probability matrix. †Bimodality means that the probability to remain in the two polar states of zero and at least 6 patents is very high, but the other probabilities are low, leading to low persistence in general. **RE means random-effects. ‡‡1993 information is missing. ††SML and SICM mean simulated maximum likelihood and stationary intertemporal covariance matrix respectively.

Table 2: Descriptive statistics

Variable	Mean	Std. Dev. Overall	Std. Dev. Between	Std. Dev. Within	Mean	Std. Dev. Overall	Std. Dev. Between	Std. Dev. Within
	Unbalanced panel				Balanced panel			
	Dependent variables							
TPP innovator	0.649				0.684			
Share of innovative sales* (if TPP=1)	0.276	0.249	0.214	0.141	0.286	0.252	0.204	0.154
	Explanatory variables							
Size†	190.649	717.912	615.813	397.759	238.355	663.468	541.293	384.144
Domestic market share‡	0.005	0.019	0.016	0.009	0.007	0.022	0.017	0.014
Part of a group	0.677				0.749			
Non R&D performer (if TPP=1)	0.202				0.164			
R&D intensity†† (if R&D>0)	0.149	1.930	1.157	1.518	0.132	1.913	0.916	1.643
Continuous R&D (if R&D>0)	0.723				0.751			
Demand pull (if TPP=1)	0.615				0.648			
Proximity to science (if TPP=1)	0.201				0.232			
Cooperation (if TPP=1)	0.346				0.383			
Subsidy (if TPP=1)	0.473				0.487			
Number of observations	7597				2352			

*A logit transformation; †ln(number of employees); ‡ln(total sales/sales of industry); ††ln(R&D/total sales) are used in the estimation.

Table 3: Transition probabilities: persistence of innovation and innovation output intensity

Innovation status		CIS					
Period t-1	Period t	CIS 2-CIS 2.5	CIS 2.5-CIS 3	CIS 3-CIS 3.5	CIS 2-CIS 2.5	CIS 2.5-CIS 3	CIS 3-CIS 3.5
		Unbalanced panel			Balanced panel		
Non-TPP	Non-TPP	59.450	63.612	80.230	56.250	66.440	81.481
	TPP	40.550	36.388	19.770	43.750	33.560	18.519
TPP	Non-TPP	19.193	22.148	33.176	14.640	20.814	27.569
	TPP	80.807	77.852	66.824	85.360	79.186	72.431

Share of innovative sales (if TPP=1)		CIS					
Period t-1	Period t	CIS 2-CIS 2.5	CIS 2.5-CIS 3	CIS 3-CIS 3.5	CIS 2-CIS 2.5	CIS 2.5-CIS 3	CIS 3-CIS 3.5
		Unbalanced panel			Balanced panel		
Below average	Below average	48.736	55.738	61.310	52.880	59.649	53.571
	Above average	51.264	44.262	38.690	47.120	40.351	46.429
Above average	Below average	18.939	23.965	25.750	14.894	21.186	23.902
	Above average	81.061	76.035	74.250	85.106	78.814	76.098

Table 4: Dynamic type 2 tobit estimates: persistence of innovation and innovation output intensity‡

Variable	Coefficient	(Std. Err.)	Coefficient	(Std. Err.)	Coefficient	(Std. Err.)	Coefficient	(Std. Err.)
	Unbalanced panel						Balanced panel	
			Individual effects				Individual effects	
	No individual effects		Exogenous initial conditions		Correlated with initial conditions		Correlated with initial conditions	
Current TPP innovator (d_{it})								
High-tech past TPP ($d_{i,t-1}$)	1.305**	(0.090)	1.289**	(0.094)	0.539**	(0.142)	0.273	(0.218)
Low-tech past TPP ($d_{i,t-1}$)	0.962**	(0.052)	0.943**	(0.055)	0.131	(0.114)	0.199	(0.157)
Lagged size (in log)	0.232**	(0.026)	0.243**	(0.028)	0.310**	(0.038)	0.475**	(0.086)
Lagged domestic market share (in log)	0.033*	(0.015)	0.037*	(0.016)	0.044*	(0.022)	0.027	(0.044)
Part of a group	0.186**	(0.048)	0.191**	(0.051)	0.236**	(0.065)	0.191	(0.117)
Intercept	-1.225**	(0.215)	-1.202**	(0.226)	-1.552**	(0.299)	-2.510**	(0.637)
Current share of innovative sales (y_{it} in logit)								
High-tech past share ($y_{i,t-1}$)	0.354**	(0.028)	0.316**	(0.032)	0.104**	(0.038)	0.132**	(0.051)
Low-tech past share ($y_{i,t-1}$)	0.280**	(0.022)	0.242**	(0.027)	0.026	(0.035)	0.077	(0.049)
Lagged size (in log)	0.114**	(0.032)	0.111**	(0.033)	0.114**	(0.034)	0.142*	(0.058)
Lagged R&D intensity (in log)	0.188**	(0.028)	0.192**	(0.029)	0.183**	(0.029)	0.185**	(0.044)
Non-R&D performer (lagged)	-1.031**	(0.191)	-1.048**	(0.193)	-0.967**	(0.194)	-0.750*	(0.309)
Continuous R&D (lagged)	-0.041	(0.089)	-0.047	(0.090)	-0.063	(0.089)	-0.048	(0.135)
Demand pull (lagged)	0.099	(0.070)	0.098	(0.070)	0.099	(0.068)	0.025	(0.103)
Proximity to science (lagged)	0.072	(0.079)	0.052	(0.080)	0.027	(0.079)	-0.076	(0.113)
Cooperation (lagged)	-0.043	(0.071)	-0.040	(0.072)	-0.033	(0.071)	0.151	(0.103)
Subsidy (lagged)	0.204**	(0.076)	0.207**	(0.076)	0.197**	(0.076)	0.118	(0.113)
Part of a group	-0.095	(0.076)	-0.101	(0.076)	-0.106	(0.076)	-0.015	(0.122)
Intercept	-0.585*	(0.245)	-0.566*	(0.250)	-0.565*	(0.249)	-0.677	(0.424)
Extra parameters								
Initial TPP (d_{i0})	-	-	-	-	1.116**	(0.143)	1.322**	(0.193)
Initial share (y_{i0})	-	-	-	-	0.220**	(0.030)	0.256**	(0.038)
σ_{a_1}	-	-	0.275**	(0.007)	0.857**	(0.070)	0.928**	(0.104)
σ_{a_2}	-	-	0.520**	(0.029)	0.882**	(0.053)	0.801**	(0.057)
σ_{ϵ_2}	1.598**	(0.070)	1.502**	(0.100)	1.335**	(0.067)	1.297**	(0.069)
$\rho_{a_1 a_2}$	-	-	0.522**	(0.077)	0.513**	(0.060)	0.582**	(0.067)
$\rho_{\epsilon_1 \epsilon_2}$	0.328**	(0.067)	0.256**	(0.081)	0.128	(0.097)	0.035	(0.134)
Number of observations			4443				1764	
Log-likelihood	-7568.809		-7565.667		-7495.589		-2936.273	

‡Note: two time dummies and three industry dummies are included in both equations.

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