The Dynamics of the Innovation Process in Dutch Manufacturing: A Panel Data Analysis

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

February 14, 2005

Abstract

This paper studies the dynamics of innovative achievement and innovation profits in Dutch using three waves of the Dutch Community Innovation Surveys pertaining to the periods 1994-1996, 1996-1998 and 1998-2000. We estimate a dynamic probit model, accounting for unobserved heterogeneity and find that past innovative achievement does not increase the probability to achieve current innovation. However, using a dynamic linear model and accounting for unobserved heterogeneity, we find a strong persistence of innovation profits among the multi-period innovators.

(VERY PRELIMINARY)

Keywords: Dynamic Panel Data, Innovation, Persistence.
JEL classification: C34, C51, O33, O38

*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 studies the persistence of innovation in Dutch manufacturing using three waves of the Community Innovation Survey (henceforth CIS), pertaining to the period 1994-2000. More specifically, we attempt to answer two questions. First, does being successful in past innovation activities increase the probability of being successful in current innovation activities? Secondly, given that one is a multi-period innovator, does past return to innovation guarantee higher current return to innovation? In other words, this study analyzes the extent to which the production of innovations is subject to “dynamic economies of scales”, or whether “success breeds success”.

The persistence of innovation is shown to be an important feature of the knowledge economy. It is very often associated with economic performance and survival of firms in the knowledge economy. Thus, various theoretical models have attempted to make predictions regarding the persistence of innovation. Four strands of the literature are considered in this study, namely: *the linear model, the financial constraints* approach*, the strategic considerations* and *the learning-by-doing* models. The *linear model* explains that persistence in innovation output coincides with persistence in innovation input, i.e. innovation is persistent only if R&D is. The *financial constraints* model states that, if a firm faces R&D funding problems, past innovation profits help to fund current innovative projects. The *strategic considerations* models states that, firms with different market power have different incentives to innovate, hence persistence of innovation is different. The idea of the last model is that, knowledge that has been used to produce past innovations can also be used to produce current innovations. The depreciation rate of innovative abilities may be very small. These theoretical models are tested empirically and help us answer our two questions.

As for empirical studies on the persistence of innovation, our analysis departs from them. Indeed, the majority of these studies use patent data to analyze persistence in innovation activities. Almost all these studies conclude alike: there is very little evidence of persistence in innovation. For instance, Geroski et al. (1997) use a duration dependence Weibull model to conclude that very few innovative UK firms are persistently innovative. Their result is robust to the type of data used in their study, i.e. patent versus major innovations data. Using a somewhat different approach, Malerba and Orsenigo (1999) also find that very few innovative firms do so persistently. A rather different result by Cefis (2003) is that, in general there is little evidence of persistence of innovation, but strong evidence among major innovators. A patent study on the persistence of innovation that is worth noting, because of a completely different result, is the one by Crépon and Duguet (1997). They estimate a dynamic count panel data using generalized method of moments (GMM) techniques to conclude that the persistence of innovation among R&D performers is very strong, as captured by the effects of past patents on current patents.

All the empirical studies previously mentioned have a common drawback: the type of the data used to measure innovation, namely patent. The term “persistence of innovation” may be misleading, since patent is rather a poor measure of innovation (see Griliches (1990) for more details). Secondly, using patent data is very demanding for persistence to be strong. Indeed, in order for a firm to be accounted for in a proper manner in the patent data set, it has to be the first to apply for a patent. In order words, when analyzing the persistence of innovation using patent data, one is unwittingly analyzing the persistence of "winning the patent race" (Duguet and Monjon (2002)). So, it is not surprising that innovation persistence is so weak, because firms are unlikely to always win the patent race, instead it happens to them once in a while. They also explain that, when using major innovations data, like in Geroski et al. (1997) one is also running into the market leader problem. Firms that implement these innovations are supposed to be the market leaders, and persistence in innovation is likely to be weak for the same reasons mentioned earlier.

Like Duguet and Monjon (2002), this study makes use of innovation survey data which consider innovation at the firm level, without mentioning the firm patent or market leadership status. But unlike Duguet and Monjon (2002), this study examines persistence in a "true state dependence" panel data framework accounting for unobserved individual-specific effects.

More specifically, we estimate two dynamic models controlling for individual effects. First, we estimate a dynamic probit model to investigate the extent to which being successful in past innovative activities affects the probability of being successful in current innovative activities, controlling for firms' characteristics. Secondly, we estimate a dynamic linear model, where the return to current innovation is explained by the return to past innovation and current innovation inputs. In this latter case, different specifications are used, according to the theoretical models underlined previously. We find that, when controlling for unobserved heterogeneity between firms, the persistence in innovative achievement vanishes. Past innovation achievement increases the probability of achieving current innovation only through unobserved effects that are correlated over time, what Heckman (1981c) calls *spurious state dependence*. However, among the multi-period innovators, even after controlling for unobserved differences between firms, we find strong persistence in innovation profits.

We explain the motivation for our analysis in Section 2, then the dynamic models we study are described in Section 3. The estimation procedures are explained in Section 4, we then briefly describe the data used to implement our models in Section 5. In Section 6 we present the results, and summarize and conclude in Section 7.

2 Motivation

The starting point of this study is the analysis by (Duguet and Monjon (2002)), or more precisely, the theoretical models mentioned in their analysis. The predictions made by these models are explained earlier in this study, and this section explains how they can be tested.

The *linear model* describes a simple relationship between firms' R&D expenditures and their innovations. Hence, consecutive innovations are explained by continuously-undertaken R&D, and are not linked to each other. Empirically speaking, the *linear model* is tested by specifying a relationship between current innovation output, and past innovation output and R&D indicators. If the *linear model* holds, once we control for R&D, past innovation should not affect current innovation, meaning that innovation is persistent only through R&D.

The *financial constraints* model gives also an explanation of the persistence of innovation. If the financial markets are imperfect, firms may face R&D funding problems. Then, an incentive to innovate is to make profits that help funding future innovative activities. To test this model empirically, we have to specify a relationship between current innovation output, and past innovation output controlling for both R&D expenditures and variables that reflect access to finance. The size of firms and received subsidies can be used to reflect access to finance. In order for this model to hold, once R&D and access to finance are controlled for, past innovation should not affect current innovation. If past innovation does influence current innovation, one should find other reasons why there is persistence in the innovation process.

The other two theoretical models that motivate the specifications of our regressions are the *strategic considerations* and the *learning-by-doing* models. In short, the first one states that, firms with different market power have different incentives to innovate. So, a market share variable should be controlled for in order to test empirically the model. If past innovation output, i.e. past innovation return affect current innovation return, after controlling for the above mentioned variables and market share, then again other reasons for the persistence in innovation should be found. The *learning-by-doing* model is the final step in specifying our regression in this study. The idea is that past innovation abilities do not immediately become obsolete, instead they are used to produce current and even future innovations. This last model motivates a relationship between current innovation output, and past innovation output controlling for R&D variables, access to finance variables, market power and variables reflecting technological opportunities and demand-pull.

Ideally, the above models can be tested using a relationship between current innovation output, and past innovation output and innovation input. They require innovation input to be available for not only firms that do innovate, but also for firms that do not innovate. For instance, one could use a dynamic probit model, where innovation output is a binary variable indicating whether or not a firm is an innovator. In this study, however, unlike Duguet and Monjon (2002), we cannot test the exact theoretical models we mention earlier. The reason is that, innovation activities, e.g. R&D expenditures, are available only for firms that do innovate. The only variables that we can use for both innovators and non-innovators are firms' characteristics like size and the type of industries they belong to.

Our approach is the following. We estimate a dynamic probit model to investigate whether dynamics is present or not, controlling for individual effects that capture unobserved heterogeneity between firms. The explanatory variables included in this model are firms' characteristics that are available for both types of firms, as explained earlier. So, the first research question is the persistence in innovative behavior, once we control for firms' characteristics and unobserved heterogeneity. We then estimate a dynamic innovation output/input regression, where the theoretical models we describe above help us specify our relationship. The measure of innovation output we use is the percentage in total sales of innovative sales, the latter being those sales that are from new or improved products at the firm level. In order to use the full set of explanatory variables as suggested by the theoretical models, a sample of multi-period innovators has to be used. In order words, we investigate the persistence of the return to innovation for firms that already have a persistent innovative behavior.

The two dynamic models are explained in the next section.

3 Dynamic models

This section describes the dynamic models used to answer our two research questions, namely the persistence in innovative behavior and of the return to innovation. The first question requires us to specify a dynamic probit equation, while the second one is answered by a dynamic linear model. The models are described as follows.

3.1 Dynamic probit

Let d_{it} be an observed binary variable with values 1 if firm i innovates at period t, and 0 otherwise. Let the decision to innovate be a latent function of past innovative achievement, the firm's characteristics and unobserved heterogeneity. Formally, the model reads:

$$d_{it} = 1\left[\rho d_{i,t-1} + \delta' \mathbf{w}_{it} + \eta_i + u_{it} > 0\right], \quad t = 1, ...T; \; i = 1, ...N, \tag{1}$$

where $d_{i,t-1}$ is a binary variable indicating whether firm i achieved at least one innovation in the past, $\mathbf{w}_{it}$ is the vector of regressors, mainly firm's i characteristics at period t, η_i and u_{it} are the individual-specific and idiosyncratic error terms, ρ is a scalar and δ a vector of parameters to be estimated. The expression in square brackets reflects the incentive to innovate, and an innovation is achieved if the incentive is high enough or crosses a certain threshold. The innovative behavior is persistent if past innovation achievement increases the probability to achieve a current innovation. Technically, a statistically significant estimated parameter ρ would explain persistence in the innovative behavior. However, a significant ρ may have another interpretation. If firms have different unobserved variables that influence the probability to innovate, and if these variables are correlated over time and are not (or improperly) controlled for, then past achievement in innovation may seem to influence current innovation achievement. This is what Heckman (1981a, 1981c) calls *pure state dependence* versus *spurious state dependence*. In our case, a pure state dependence would correspond to a *true* persistence in the innovative behavior, while a spurious state dependence would correspond to a persistence in innovative behavior through a correlation over time of unobserved variables. This is a bit similar to the idea of persistence in the theoretical *linear model* discussed earlier. The only exception is that, the theoretical *linear model* predicts persistence in innovation through the persistence or correlation over time of *observed* variables, namely, the ones pertaining to R&D, while spurious state dependence predicts persistence in innovation through the correlation over time of *unobserved* variables.

So, unlike Duguet and Monjon (2002), we attempt to account for both dynamic and unobserved heterogeneity in Eq. (1). We now turn to the specification of the dynamic linear model.

3.2 Dynamic linear model

Let y_{it} be the innovation output, as measured by the share of innovative sales, of firm i at period t. Investigating the effect of past return on current return to innovation leads us to write:

$$y_{it} = \gamma y_{i,t-1} + \beta' \mathbf{x}_{it} + \alpha_i + \varepsilon_{it}, \quad t = 1, ...T; \; i = 1, ...N, \tag{2}$$

where γ and β are respectively a scalar and a vector of parameters to be estimated, $\mathbf{x}_{it}$ a vector of regressors that are chosen according to the theoretical models we discuss earlier, and α_i and ε_{it} are the individual-specific and idiosyncratic error terms. This equation is estimated conditional on the fact that firm i is a multi-period innovator, for the reasons we explained earlier in the study. Hence, Eq. (2) studies the persistence in the innovation return for multi-period innovators. This approach is similar to the analysis by Crépon and Duguet (1997) who study the persistence in patenting for R&D performers. Patent is replaced in this study by the share of innovative sales.

We now discuss the estimation procedures of our dynamic models.

4 Estimation

This section addresses the issue of estimating dynamic models accounting for unobserved heterogeneity. We start by discussing this issue for the dynamic probit and then turn to the dynamic linear model.

4.1 Dynamic probit

When estimating discrete choice panel data models, an often-encountered difficulty is the individual effects term. In static fixed-effects logit models, a way to handle the "incidental parameters problem", is to maximize a likelihood function conditional on a sufficient statistic which is shown to be $s_i = \overline{d}_i = \frac{1}{T} \sum_{t=1}^{T} d_{it}$. However, the conditional probability of d_{i1},...,d_{iT} is degenerate if $s_i = 0$ or $s_i = 1$, meaning that individuals that remain in the same state over time are discarded when a fixed-effects logit model is used. This model is not applicable to our case, because we would model the probability of switching from the non-innovative state to the innovative state and vice-versa, while we would ignore firms that always or never innovate. Another reason why the fixed-effects logit model is not used in this study, is because it does not allow us to control for technological opportunities via industry dummies, which are constant or show very little variation over time. So, a random-effects approach is considered in this study, the specification of which is described later.

Another difficulty encountered in discrete choice panel data models, when dynamic is involved, is the "initial conditions" problem. In the literature, two assumptions are often made about the initial conditions: 1) the initial conditions are exogenous, or 2) the process is assumed to be in equilibrium. Neither assumption is satisfactory (Hsiao (2003)), and a proper way to treat the initial conditions has to be found. Heckman (1981b) and Wooldridge (2002) describe two approaches to handle the problem. We follow their approach.

4.1.1 The Heckman approach

The random-effects approach that we follow in this study is specified as follows: η_i is normally distributed with mean 0 and standard deviation σ_η, u_{it} follows a standard normal distribution for all t.

Heckman (1981b) suggests that the initial conditions be approximated by a probit model

$$d_{i0} = 1[\theta' \mathbf{w}_{i0} + \lambda \eta_i + u_{i0} > 0]. \tag{3}$$

$\mathbf{w}_{i0}$ is the vector of regressors that affect the probability to innovate at the "initial" period, u_{i0} the error term at the initial period, θ and λ are respectively the vector of parameters and a scalar to be estimated.[1] The latter captures the dependence between the initial values and the individual effects. If we denote $\varsigma_{i0} = \lambda\eta_i + u_{i0}$ the error term in Eq. (3), the correlation term between ς_{i0} and η_i denoted $corr(\varsigma_{i0}, \eta_i)$ can be written as $corr(\varsigma_{i0}, \eta_i) = \frac{\lambda\sigma_\eta}{\sigma_{\varsigma_{i0}}}$ (Heckman (1981b)), where $\sigma_{\varsigma_{i0}}$ is the standard deviation of the error term of the initial probit equation.

Once the initial conditions problem is handled, the model, as specified in Eqs. (1) and (3), can be estimated by maximum likelihood. The contribution of one individual to the likelihood is written as

$$LH_i(d_{i0}, d_{i1}..., d_{iT}|d_{i,t-1,}\mathbf{w}_{it}) = \int_{-\infty}^{\infty} \prod_{t=1}^{T} \Phi\left[(2d_{it}-1)\left(\rho d_{i,t-1} + \delta'\mathbf{w}_{it} + \eta_i\right)\right] \times \Phi\left[(2d_{i0}-1)\left(\theta'\mathbf{w}_{i0} + \lambda\eta_i\right)\right] \frac{1}{\sigma_\eta}\phi\left(\frac{\eta_i}{\sigma_\eta}\right) d\eta_i. \quad (4)$$

4.1.2 The Wooldridge approach

Another approach to handle the initial conditions problem is suggested by Wooldridge (2002). Instead of specifying or approximating (Heckman (1981b)) the distribution of the initial outcome given the individual effects, Wooldridge suggests specifying the distribution of the individual effects given the initial outcome. So, the joint distribution of all outcomes in Eq. (4) becomes a distribution of the outcomes starting form $t = 1$, given the initial outcome $d_{i0.}$ Wooldridge's suggestion is to write the individual effects as

$$\eta_i = b_0 + b_1 d_{i0} + \mathbf{b}_2'\mathbf{z}_i + a_i, \quad (5)$$

with $a_i \rightarrow N(0, \sigma_a)$, $\mathbf{z}_i' = (\mathbf{z}_{i1}', ...\mathbf{z}_{iT}')$, and b_0, b_1 and $\mathbf{b}_2'$ are respectively scalars and a vector of parameters to be estimated. The distribution of the individual effects conditional on the initial outcome is then normal with mean $b_0 + b_1 d_{i0} + \mathbf{b}_2'\mathbf{z}_i$ and standard deviation σ_a. The likelihood function for one individual can then be written as

$$LW_i(d_{i1}..., d_{iT}|d_{i0}, d_{i,t-1,}\mathbf{w}_{it}) = \int_{-\infty}^{\infty} \prod_{t=1}^{T} \Phi[(2d_{it}-1)\left(\rho d_{i,t-1} + \delta'\mathbf{w}_{it} + b_0 + b_1 d_{i0} + \mathbf{b}_2'\mathbf{z}_i + a_i\right)] \frac{1}{\sigma_a}\phi\left(\frac{a_i}{\sigma_a}\right) da_i. \quad (6)$$

While the Heckman approach is shown to perform quite well in Monte Carlo studies, the Wooldridge approach is more computationally attractive.[2] We now turn to the estimation of the dynamic linear model.

[1] In order for θ' to be identified, at least one regressor in w_{i0} should not be in w_{it}.

[2] While the likelihood expression in Eq. (6) has the same struture as the likelihood of the standard random-effects probit model, the initial out come is not treated as independent of the individual effects. The dependence is captured by the scalar b_1.

4.2 Dynamic linear model

To estimate the dynamic linear model, we adopt a random-effects approach, as described in Anderson and Hsiao (1981, 1982). One reason is the one mentioned previously in the study, i.e., it allows us to estimate the effects of technological opportunities as captured by industry dummies, while a GMM approach would difference out these effects. We estimate the model by maximum likelihood and make the following assumptions: α_i and ε_{it} are normally distributed and

$$\begin{aligned} E\alpha_i &= E\varepsilon_{it} = 0, \\ E\alpha_i \mathbf{x}'_{it} &= \mathbf{0}', \ E\alpha_i\varepsilon_{it} = 0 \\ E\alpha_i\alpha_j &= \begin{cases} \sigma_\alpha^2 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases} \\ E\varepsilon_{it}\varepsilon_{js} &= \begin{cases} \sigma_\varepsilon^2 & \text{if } i = j, \, t = s \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

As in the dynamic probit, the initial outcome is also of great issue in the dynamic linear model. The idea is the same as before, we have to specify a relationship between the initial outcome and the individual effects. Following Anderson and Hsiao (1981, 1982), the initial outcome is assumed to be random and correlated with the individual effects. The intuition is that, the initial outcome affects all future outcomes through its correlation with the individual effects. The initial outcome is then assumed to be normally distributed with mean $\mu_{y_{i0}}$ and standard deviation $\sigma_{y_{i0}}$. Furthermore, the covariance between the initial outcome and the individual effects is denoted by $\tau\sigma^2_{y_{i0}}$, and the parameters $\mu_{y_{i0}}$, $\sigma_{y_{i0}}$ and τ are to be estimated. The likelihood function, taking into account the joint distribution of all outcomes starting from $t = 0$, is given by

$$\begin{aligned} L = {} & (2\pi)^{-\frac{NT}{2}} (\sigma_\varepsilon^2)^{-\frac{N(T-1)}{2}} (\sigma_\varepsilon^2 + Tc)^{-\frac{N}{2}} \\ & \times \exp\left\{ -\frac{1}{2\sigma_\varepsilon^2} \sum_{i=1}^{N} \sum_{t=1}^{T} [y_{it} - \gamma y_{i,t-1} - \beta' \mathbf{x}_{it} - \tau(y_{i0} - \mu_{y_{i0}})]^2 \right. \\ & \left. + \frac{c}{2\sigma_\varepsilon^2(\sigma_\varepsilon^2 + Tc)} \sum_{i=1}^{N} \left\{ \sum_{t=1}^{T} [y_{it} - \gamma y_{i,t-1} - \beta' \mathbf{x}_{it} - \tau(y_{i0} - \mu_{y_{i0}})] \right\}^2 \right\} \\ & \times (2\pi)^{-\frac{N}{2}} (\sigma^2_{y_{i0}})^{-\frac{N}{2}} \exp\left\{ -\frac{1}{2\sigma^2_{y_{i0}}} \sum_{i=1}^{N} (y_{i0} - \mu_{y_{i0}})^2 \right\}, \end{aligned} \tag{7}$$

where $c = \sigma_\alpha^2 - \tau^2\sigma^2_{y_{i0}}$. The likelihood function in Eq.(7) is the product of a likelihood taking account of the joint density of the outcomes starting from $t = 1$ conditional on the initial outcome, and the likelihood of the marginal density of the initial outcome.

We explain in the next section the data used to implement our models.

5 Data

In order to implement our models, three waves of the Dutch CIS, pertaining to the periods 1994-1996, 1996-1998 and 1998-2000, are used. The sample consists of Dutch manufacturing firms, with at least 10 employees , that existed in 1994 and survived until 2000. We form our panel using firms that took part in the three innovation surveys (balanced panel).

The variables used in our regressions are described as follows. The dependent variable used in the probit model indicates whether or not a firm is a technological product and/or process innovator; and the one used in the regression is the percentage in total sales of innovative sales. In order to make the latter variable lie within the real number interval, a logit transformation is considered. As explanatory variables of the probit model, firms' characteristics like size and the industry they belong to are included. The variable size is measured by the number of employees, in natural logarithm, and the industry in which a firm operates is given by the Dutch standard industrial classification (SBI 1993). As for the regression's explanatory variables, they are included according to the theoretical models discussed above. Besides size and industry dummies, three R&D variables, an access-to-finance variable, market share, and two variables used as proxies for demand pull and technology push are used. More details on data collection, the construction of the variables and descriptive statistics for each industry in the three innovation surveys can be found in Raymond et al. (2004).

The empirical results are discussed in the next section.

6 Empirical results

We start by discussing the persistence in innovation achievement (dynamic probit) and then turn to the persistence in the return to innovation (dynamic linear model).

6.1 Persistence in innovation achievement

An interesting point that is worth noting is that both the Heckman and the Wooldridge approaches show similar results in terms of the sign and significance of the variables. Innovation achievement of a firm at a certain period of time depends strongly and significantly on its size, as measured by the number of employees. This confirms the result of a positive and significant effect of size on the probability to innovate in cross-sectional studies (see Table 11 of Raymond et al. (2004)). Market share, which reflects the differences in innovative incentive of firms, according to the *strategic considerations* theory, does not seem to influence the probability to innovate. But, we could imagine that larger firms have larger market shares, so that the two variables are correlated, and that the significant effect of one variable lessens the significance of the other variable. Other things being equal, the probability to achieve an innovation is higher in the chemical industry. Tables (1) and (2) also show that innovate at the "initial" period influences strongly, positively and significantly the probability to innovate at a certain period of time. This influence is either direct, as shown in Table (1), or indirect through "initial" regressors or through the correlation term between the initial outcome and the individual-specific term (Table (2)).

The most interesting result shown by Tables (1) and (2) is the following. Once we account for unobserved heterogeneity, as captured by the standard deviations σ_η and σ_a the persistence in innovation achievement vanishes.[3] This result contradicts the result by Duguet and Monjon (2002) who claim that, the lack of persistence in innovation, that is usually found in empirical studies, is due to the type of the data used to measure innovation. They indeed find a strong evidence of innovation, in the sense that past innovation achievement strongly influences current innovation achievement. However, they do not control for unobserved differences that may exist between firms. The result of strong persistence in innovation that they find in their study, is in fact a strong persistence in unobserved differences between firms, what Heckman (1981c) calls spurious state dependence. Indeed, we find that there is significant unobserved differences between firms, as shown by the highly-significant standard deviations σ_η and σ_a. When these differences are correlated over time and are not accounted for, the effect of past innovation achievement is overstated (which is probably the case in Duguet and Monjon (2002)). A chi-square test, (which is not reported in the tables) of no correlation of unobserved differences, i.e. $\frac{\sigma_\eta^2}{\sigma_\eta^2+1} = 0$, or $\frac{\sigma_a^2}{\sigma_a^2+1} = 0$ is clearly rejected.[4]

We now discuss the results on the persistence of the return to innovation for multi-period innovators.

6.2 Persistence of the return to innovation

This section presents the empirical results when implementing the theoretical models discussed in Section (2). We start by testing the *linear model*, which states that, if there is persistence in innovation, it is due to persistence in R&D activities. The *theoretical linear model*, as considered in this study, is slightly different from the *linear model* considered in Duguet and Monjon (2002). In our analysis, the question is that, given that a firm is a multi-period innovator, what guarantees the firm higher return to innovation as time goes by, especially does its past return to innovation affect its current return? In this context, the *linear model* holds if the persistence in innovation return is due to the persistence R&D activities. Table (3) shows that, *ceteris paribus*, multi-period innovators that do perform R&D, and that do so on a continuous basis have higher return to innovation. The more money they spend in R&D activities, the higher their return to innovation, other things being equal. However, the effect of past return on current return remains strongly significant, indicating that there may be other factors influencing the persistence of the return to innovation. The *theoretical linear model* is then only partially supported.

We then include additional regressors according to the *financial constraints* (Table (4) and the *strategic considerations* (Table (5) models. Past innovation return remains statistically strongly significant, while neither variable of the former model nor variable of the latter is significant. The persistence in the

[3]Actually, we have also implemented a dynamic probit model without accounting for unobserved heterogeneity. We found then a strongly significant and positive effect of past innovation achievement on current innovation achievement.

[4]In fact, we could have estimated two more general models, namely a first-order Markov process, or a state-dependent model with stationary intertemporal covariance matrix (SICM) (Heckman 1981a, 1981c). But, the error-component formulation, which is a special case of the SICM model rejects already the "true" persistence of innovation. We then expect the other two models to reject the "true state dependence as well.

innovation return is not explained by easier access to finance or higher market power.

Finally, two additional variables, namely demand pull and technology push are included leading us to a *learning-by-doing* model. The effect of past innovation return on current innovation return remains highly significant. Furthermore, demand pull has a highly significant effect on the return to innovation, the effect of technology push is lesser.

Table (6) suggests that, among the multi-period innovators, the return to innovation is strongly significant, even after controlling for unobserved heterogeneity, as captured by σ_α. There are various reasons for this. First, this persistence is partly explained by a persistence in R&D activities. However, this is not the only reason, as suggested by the *linear model*, a *learning-by-doing* explanation can be found as well. Indeed, the effect of R&D activities on the persistence of innovation return lessens when demand pull and technology push are included in the model. So, a second reason, learning-by-doing, is that past knowledge that allowed firms to innovate and generate past profits is used to produce current innovation and make current profits. We can even say that, an innovation that generated profit in the past, say 3 years ago, is not immediately obsolete and continues to generate profit. A third explanation in the persistence of return to innovation is that of demand pull, which in fact is closely related to the second explanation. In our analysis, demand pull is constructed as a variable that measures the inclination of a firm to 'open up new markets', 'extend product range' or 'replace products phased out'. Past innovative products, through the profits they generated, can "show the way" how to achieve the objectives mentioned above and generate new profits.

7 Conclusion

This paper shows that, when estimating the dynamics of innovative behavior in Dutch manufacturing and accounting for unobserved heterogeneity, being successful in the past innovation does not increase the probability to innovate in the future. There is no *true state dependence* in the innovative behavior in Dutch manufacturing. The only persistence in innovative behavior that occurs is through unobserved effects that are correlated over time and that affect the probability to innovate. There is *spurious state dependence* in the innovative behavior of firms in Dutch manufacturing.

However, we find evidence of strong persistence in the innovative return to innovation among multi-period innovators. This persistence is partly explained by persistence in R&D activities, but also by learning-by-doing and demand pull. We find no evidence of the effects of access-to-finance and *strategic constraints* variable, as captured by size, subsidies received and market share.

An explanation is as follows. In another cross-sectional study (Raymond et al (2004)), we found that, other things being equal, smaller firms have higher return to innovation. But, small firms are less likely to survive for longer periods, or to take part in all the innovation surveys than larger firms. The effect of size may be altered by this phenomenon. So, future research should investigate whether our results are sensitive to attrition.

Because of the way the CIS is designed, we have estimated our dynamic linear model conditional on being a multi-period innovator. The sample is somewhat reduced compared to a situation where non multi-period innovators are included in the analysis. We could link the two equations by allowing that the error terms in the probit and in the regression equations are correlated and estimate a tobit type II, according to the terminology of Amemiya (1985), in a dynamic panel data framework. The studies by Kyriazidou (1997, 2001) could be of guidance.

References

[1] Amemiya, T. (1985), *Advanced Econometrics*, Harvard University Press, Cambridge.

[2] Anderson, T., and C. Hsiao (1981), "Estimation of Dynamic Models with Error Components." *Journal of the American Statistical Association*, 76, 598-606.

[3] Anderson, T. W. and C. Hsiao (1982), "Formulation and Estimation of Dynamic Models Using Panel Data," *Journal of Econometrics* 18, 67-82.

[4] Cefis, E. (2003), "Is there persistence in innovative activities", *International Journal of Industrial Organization, 21, 489-515.*

[5] Crépon, B., E. Duguet (1997), "Estimating the Innovation Function from Patent Numbers: GMM on Count Panel Data", *Journal of Applied Econometrics*, 12, 243-263.

[6] Duguet, E. and S. Monjon (2002), "Creative Destruction and the Innovative Core: Is Innovation Persistent at the firm level?" *mimeo.*

[7] Geroski, P., J. Van Reenen and C. Walters (1997), "How persistently do firms innovate?", *Research Policy*, 26, 33-48.

[8] Griliches, Z. (1990), "Patent Statistics as Economic Indicators: A Survey" *Journal of Economic Literature*, 28, 1661-1707.

[9] Heckman, J. J. (1981a), "Statistical Models for Discrete Panel Data", in *Structural Analysis of Discrete Data with Econometric Applications,* edited by C. F. Manski and D. McFadden, 114-178. Cambridge, MA: MIT Press.

[10] Heckman, J. J. (1981b), "The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time-Discrete Data Stochastic Process", in *Structural Analysis of Discrete Data with Econometric Applications,* edited by C. F. Manski and D. McFadden, 114-178. Cambridge, MA: MIT Press.

[11] Heckman, J. J. (1981c), "Heterogeneity and State Dependence", in *Studies in Labor Markets,* edited by S. Rosen, 91-139. University of Chicago Press.

[12] Hsiao, C. (2003), Analysis of Panel Data, Cambridge University Press, 2nd Ed..

[13] Kyriazidou E. (1997), "Estimation of a Panel Data Sample Selection Model", Econometrica, 65, 1335-1364.

[14] Kyriazidou E. (2001), "Estimation of Dynamic Panel Data Sample Selection Models", *Review of Economic Studies*, 68, 543-572.

[15] Malerba, F. and L. Orsenigo (1999), "Technological entry, exit and survival: an empirical analysis of patent data", *Research Policy*, 28, 643-660.

[16] Raymond, W., P. Mohnen, F. Palm and S. S. v. d. Loeff (2004), "An Empirically-Based Taxonomy of Dutch Manufacturing: Innovation Policy Implications, *CESifo Working Paper*.

[17] Wooldridge, J. M., (2002), "Simple Solutions to the Initial Conditions Problem in Dynamic, Nonlinear Panel Data Models with Unobserved Heterogeneity", *Working Paper*.

A Appendix

Table 1: Dynamic random-effects probit estimates: the Wooldridge approach

Variable	Coefficient	(Std. Err.)
Current innovation achievement (d_{it})		
Past innovation achievement ($d_{i,t-1}$)	0.11	(0.20)
Initial innovation achievement (d_{i0})	1.23**	(0.25)
Size	0.38**	(0.11)
Market share	-0.01	(0.08)
Chemicals	0.71**	(0.22)
Electrical	0.48*	(0.22)
Machinery & Equipment	0.52**	(0.18)
Vehicle	0.29	(0.23)
Plastic	0.55*	(0.28)
Products not elsewhere classified	0.40	(0.25)
Food and Tobacco	0.15	(0.19)
Metallic	0.32*	(0.16)
Non-metallic	0.24	(0.28)
Textile	0.02	(0.30)
Intercept	-2.32*	(1.01)
σ_a	0.77**	(0.17)
Number of firms	1722	
Log-likelihood	-848.82	
$\chi^2_{(14)}$	183.87	
Significance levels : $\dagger$: 10% $*$: 5% $**$: 1%		

Table 2: Dynamic random-effects probit estimates: the Heckman approach

Variable	Coefficient	(Std. Err.)
Current innovation achievement (d_{it})		
Past innovation achievement ($d_{i,t-1}$)	0.10	(0.20)
Size	0.44^{**}	(0.12)
Market share	0.00	(0.08)
Chemicals	0.95^{**}	(0.25)
Electrical	0.77^{**}	(0.25)
Machinery & Equipment	0.77^{**}	(0.21)
Vehicle	0.29	(0.23)
Plastic	0.84^{**}	(0.31)
Products not elsewhere classified	0.39	(0.25)
Food and Tobacco	0.26	(0.20)
Metallic	0.36^{*}	(0.17)
Non-metallic	0.23	(0.28)
Textile	0.01	(0.30)
Intercept	$-1.75^{\dagger}$	(0.99)
Initial innovation achievement (d_{i0})		
Size at period zero	0.26^{**}	(0.08)
Chemicals	0.76^{**}	(0.29)
Electrical	1.07^{**}	(0.33)
Machinery & Equipment	0.91^{**}	(0.27)
Plastic	1.03^{**}	(0.38)
Food and Tobacco	0.35	(0.26)
Metallic	0.15	(0.21)
Intercept	-0.53	(0.36)
σ_η	0.99^{**}	(0.20)
$corr(\varsigma_{i0}, \eta_i)$	0.74^{**}	(0.04)
Number of firms	1722	
Log-likelihood	-1323.89	
$\chi^2_{(13)}$	94.51	

Significance levels : $\dagger$: 10% $*$: 5% $**$: 1%

Table 3: Dynamic linear random-effects estimates: the *linear model*

Variable	Coefficient	(Std. Err.)
Current share of innovative sales (y_{it})		
Past share of innovative sales ($y_{i,t-1}$)	0.17**	(0.06)
Non-R&D performers	-2.07**	(0.75)
R&D intensities	0.28*	(0.12)
Performing continuous R&D	0.93*	(0.37)
Chemicals	1.03†	(0.56)
Electrical	0.27	(0.60)
Machinery & Equipment	1.02†	(0.54)
Plastic	0.24	(0.64)
Vehicle	1.31†	(0.67)
Food and Tobacco	0.35	(0.59)
Metallic	0.08	(0.51)
Non-metallic	-0.41	(0.78)
Products not elsewhere classified	-0.03	(0.68)
Textile	0.17	(0.89)
Intercept	-0.86	(0.81)
σ_α	1.36**	(0.37)
σ_ε	3.48**	(0.14)
$\mu_{y_{i0}}$	-0.93**	(0.20)
$\sigma_{y_{i0}}$	4.32**	(0.14)
$corr(y_{i0}, \alpha_i)$	0.37**	(0.11)
Number of firms	1347	
Log-likelihood	-3741.38	
$\chi^2_{(14)}$	87.78	

Significance levels : $\dagger$: 10% $\quad *$: 5% $\quad **$: 1%

Table 4: Dynamic linear random-effects estimates: the *financial constraints model*

Variable	Coefficient	(Std. Err.)
Current share of innovative sales (y_{it})		
Past share of innovative sales ($y_{i,t-1}$)	0.17**	(0.06)
Non-R&D performers	-1.84*	(0.75)
R&D intensities	0.25*	(0.12)
Performing continuous R&D	0.78*	(0.38)
Size	0.14	(0.13)
Subsidies	0.47	(0.30)
Chemicals	0.87	(0.56)
Electrical	0.17	(0.60)
Machinery & Equipment	0.91†	(0.54)
Plastic	0.11	(0.64)
Vehicle	1.18†	(0.67)
Food and Tobacco	0.22	(0.59)
Metallic	-0.01	(0.51)
Non-metallic	-0.52	(0.78)
Products not elsewhere classified	-0.10	(0.68)
Textile	0.22	(0.89)
Intercept	-1.80†	(1.02)
σ_α	1.33**	(0.38)
σ_ε	3.48**	(0.14)
$\mu_{y_{i0}}$	-0.93**	(0.20)
$\sigma_{y_{i0}}$	4.32**	(0.14)
$corr(y_{i0}, \alpha_i)$	0.36**	(0.12)
Number of firms	1347	
Log-likelihood	-3739.36	
$\chi^2_{(16)}$	93.25	

Significance levels : † : 10% $*$: 5% $**$: 1%

Table 5: Dynamic linear random-effects estimates: the *strategic considerations model*

Variable	Coefficient	(Std. Err.)
Current share of innovative sales (y_{it})		
Past share of innovative sales ($y_{i,t-1}$)	0.17**	(0.06)
Non-R&D performers	-2.01**	(0.77)
R&D intensities	0.29*	(0.12)
Performing continuous R&D	0.75*	(0.38)
Size	-0.24	(0.32)
Subsidies	0.45	(0.30)
Market share	0.34	(0.26)
Chemicals	0.99†	(0.57)
Electrical	0.00	(0.62)
Machinery & Equipment	0.77	(0.55)
Plastic	-0.39	(0.75)
Vehicle	1.06	(0.68)
Food and Tobacco	0.37	(0.60)
Metallic	-0.02	(0.51)
Non-metallic	-0.97	(0.85)
Products not elsewhere classified	-0.55	(0.77)
Textile	-0.44	(1.02)
Intercept	2.34	(3.30)
σ_α	1.34**	(0.39)
σ_ε	3.47**	(0.14)
$\mu_{y_{i0}}$	-0.93**	(0.20)
$\sigma_{y_{i0}}$	4.32**	(0.14)
$corr(y_{i0}, \alpha_i)$	0.36**	(0.11)
Number of firms	1347	
Log-likelihood	-3738.49	
$\chi^2_{(17)}$	95.38	
Significance levels :	† : 10% ∗ : 5% ∗∗ : 1%	

Table 6: Dynamic linear random-effect estimates: the *learning-by-doing model*

Variable	Coefficient	(Std. Err.)
Current share of innovative sales (y_{it})		
Past share of innovative sales ($y_{i,t-1}$)	0.18**	(0.06)
Non-R&D performers	-1.59*	(0.77)
R&D intensities	0.24*	(0.12)
Performing continuous R&D	0.72†	(0.38)
Size	-0.30	(0.31)
Subsidies	0.34	(0.30)
Market share	0.34	(0.25)
Demand pull	1.00**	(0.28)
Technology push	0.58†	(0.31)
Chemicals	0.93†	(0.56)
Electrical	-0.18	(0.61)
Machinery & Equipment	0.73	(0.54)
Plastic	-0.45	(0.74)
Vehicle	0.86	(0.67)
Food and Tobacco	0.24	(0.59)
Metallic	-0.03	(0.50)
Non-metallic	-1.21	(0.84)
Products not elsewhere classified	-0.64	(0.75)
Textile	-0.45	(1.01)
Intercept	1.81	(3.27)
σ_α	1.21**	(0.44)
σ_ε	3.47**	(0.15)
$\mu_{y_{i0}}$	-0.93**	(0.20)
$\sigma_{y_{i0}}$	4.32**	(0.14)
$corr(y_{i0}, \alpha_i)$	0.35**	(0.13)
Number of firms	1347	
Log-likelihood	-3729.76	
$\chi^2_{(19)}$	115.97	

Significance levels : † : 10% $*$: 5% $**$: 1%