TILBURG ◆ UNIVERSITY

TILEC

# TILEC Discussion Paper

DP 2007-008

## The Innovation Threshold

By

Erik Brouwer, Tom Poot, Kees van Montfort

February 2007

ISSN 1572-4042

# The Innovation Threshold

by
Erik Brouwer, Tilburg University, The Netherlands
Tom Poot, Utrecht University, The Netherlands
Kees van Montfort, Free University Amsterdam, Nyenrode Business University, The Netherlands[1]

*Abstract*

In this paper, we propose an economic model to analyse the sales out of new products. This model accounts for the fact that even among firms for which R&D is a permanent activity, a fraction of firms does not have sales of innovative products during a two-year observation period. We propose a model in which the fixed costs of introduction is a major concern in the decision making process. In a structural model we estimate the fixed costs of the market introduction of new products and explain subsequent sales of innovative products. We examine a indicator of innovative output, i.e. sales of products 'new to the firm'. We estimate fixed costs thresholds using data from the Dutch part of the Community Innovation Survey (CIS) from 1998. R&D intensity, competition and market structure have a positive impact on sales of new products. The most important factors to decrease the fixed costs threshold of introduction are product related R&D investments, R&D subsidies and knowledge spillovers.

Keywords: Innovation, Product R&D, Threshold model
JEL Classification: C51 (Model construction and estimation), L23 (Organization of production), O30 (Technological change, research and development)

## 1 Introduction

In this paper we present a new model to provide insight in the innovation process. This model accounts for the fact that even among firms for which R&D is a permanent activity, a fraction of firms does not have sales of innovative products during a two-year observation period. We propose a model in which the fixed costs of introduction is a major concern in the decision making process. In this study we analyse an indicator of a firm’s innovative output, i.e. sales per employee of products that are ‘new to the firm’. This indicator has been analysed in Brouwer and Kleinknecht (1996) first. They noted that, during the observation year, many firms had no sales out of new products and proposed an ad hoc model to deal with the observations with zero sales. A drawback of their procedure is that the parameter estimates do not have a clear interpretation.

[1] Corresponding author.
We would like to thank Geert Ridder and Alfred Kleinknecht for their valuable comments and discussion.

In this study we start from the assumption that, in order to introduce a product innovation, the firm must incur fixed costs. At the time of considering a new product introduction, the firm will compare fixed costs with expected (net) revenues during the economic lifetime of the product, and only if revenues exceed fixed costs, the firm will introduce the product. We show that this simple theoretical model leads to a specific censored regression model, the stochastic threshold model. Moreover, from theoretical considerations one can argue that variables that affect fixed costs do not have a direct effect on the level of sales. This suffices to identify the parameters of the threshold model. Although, without further assumptions, we cannot estimate fixed costs directly, we can identify determinants of the threshold and of the sales of innovative products. This is an improvement on the estimates obtained in Brouwer and Kleinknecht (1996) who did not distinguished between these two equations. Their estimates are reduced-form estimates, while we present structural estimates that correspond to a structural economic model, and allow for an interpretation in the context of that model.

Our paper is organised as follows. In section 2 we introduce the theoretical model and describe the corresponding econometric model. Section 3 discusses our data. Section 4 presents our hypotheses. In section 5 we present and discuss the outcomes of our estimates, and conclusions are drawn in section 6.

## 2 Sales of innovative products with fixed costs of introduction

In this paper we deal with a theoretical and empirical model based on Brouwer (1997). Some investments have the character of irreversible sunk costs, which explains the observation that the resale or scrap value of certain investments (e.g. some specialised pieces of machinery) is much lower than the purchasing price. The introduction of an innovative product that is new to the firm is also imperfectly reversible and involves fixed costs. Such fixed costs can be related to investment in specific knowledge, acquisition of patent rights, training, market research, advertising, or purchase of specialised equipment. In this section we develop a theoretical model that takes into account these fixed costs of introduction. This model implies that the innovative product is introduced if (expected) sales exceed a certain threshold. We then specify a statistical model that corresponds to this economic model, and we show that the restrictions that can be derived from the economic model are sufficient to identify this threshold, which we assume to be firm specific.

### 2.1 The theoretical model

Besides fixed costs of introduction, there may be other reasons of why an R&D performing firm does not have sales of innovative products in a particular year. The product may be still under development or the firm may have realised only *process* innovations that were asked for in a yes/no question (without assessing sales related to it). In our analysis, we concentrate on firms that engage *permanently* (other than occasionally) in R&D and that have been developing new products in the recent past. Wider definitions of sample firms have also tentatively been used in our estimates but will not be documented in detail, since these analyses did not lead to substantially different outcomes. We assume that our R&D performing firms must decide in any year whether to introduce some innovative product.

First, we introduce the variables used in the model.

| | |
|---|---|
| $s$ | = R&D spending; |
| $y$ | = annual sales of the innovative product; |
| $\mu$ | = expected annual sales of the innovative product; |
| $c$ | = fixed cost of introduction of the innovative product; |
| T | = economic lifespan of the innovative product; |
| r | = discount rate; |
| γ | = mark-up of the price over variable cost per unit. |

Without loss of generality we assume that all R&D spending is done at time 0, just before the introduction decision. At that time the expected discounted return from the investment *s* is

$$(1) \qquad \int_0^T \gamma.\mu(\mathrm{s}).e^{-rt}.dt - \mathrm{s} - \mathrm{c}$$

We assume that R&D spending *s* has a positive effect on the expected sales of the innovative product. The first-order condition for a profit maximum is

$$(2) \qquad \frac{\gamma}{r}.\left(1-e^{-rT}\right).\mu'(s) = 1$$

If we assume that $\mu''(s) < 0$ and $\mu'(s) > 0$, then the first-order equation has a finite solution which may be 0. We denote this optimal R&D investment by $s^{opt}$. The optimal level of R&D spending is independent of the fixed cost of introduction *c*. Furthermore, R&D spending rises with γ, which is an index of competition. The stronger competition, the lower is the optimal R&D spending. Unfortunately we do not have an index of competition. Instead we use a Herfindahlindex as a proxy of competition.[2] The optimal level rises also with T, and it falls with r. We unfortunately have no indicator of the economic lifespan of the innovative product at our

2 Strictly speaking a Herfindahlindex of concentration is not an equivalent of an index of competition. However in concord with the Structure-Conduct-Performance framework a concentration index can be regarded as a proxy of competition (Church & Ware, 2000).

disposal. This may seem an important omission in order to identify the proposed model. However from a an entrepreneurial point of view future sales are unknown and the decision to introduce a new product onto the market will depend on expected sales, which will depend on demand i.e. consumer preferences, and product substitution and imitation, in short competition, which is already incorporated in the model. We will elaborate on supply, demand and market conditions later on.

Note that $\mu(s)$ is the expected sales at the time that the level of R&D spending is determined. In this model these expenditures on R&D are called the current R&D spending which are the costs the entrepreneur is expected to recover from sales to distinguish from the R&D spending in the past which can be considered as sunk cost, which by definition have no relevance in the decision making process whether or not to introduce new products onto the market. At the moment that the decision whether to introduce the innovative product is taken, the expected sales differ from their previously expected value by a prediction error ν and are equal to

(3) $$y = \mu(s^{opt}) + \nu$$

The innovation is introduced if and only if the expected discounted return exceeds the fixed cost of introduction

(4) $$\frac{\gamma.\left(\mu\left(s^{opt}\right) + \nu\right).\left(1 - e^{-r.T}\right)}{r} > c$$

Equation (3) and inequality (4) specify a threshold regression model: if the expected revenues exceed the threshold costs (in inequality (4)) then sales are given in equation (3).

This simple theoretical model implies that the optimal level of R&D spending and hence the expected sales of the new product, is independent of the cost of introduction. This does not mean that the R&D effort is independent of the costs of introduction. If the expected costs of introduction are high, the project may be unprofitable, i.e. the expected discounted return in (1) may be negative, and the project will not be implemented. However, given that a project is profitable, the optimal spending is independent of the fixed costs of introduction. This result is important for the specification of the threshold model, because it justifies the exclusion of variables that affect the fixed cost of introduction from the sales equation. It should be stressed that in our empirical model R&D effort is an explanatory and not a dependent variable, i.e. we model sales of innovative products given (past) R&D effort.

### 2.2 The econometric model

For the specification of the econometric model that corresponds to equations (3) and (4), it is convenient to introduce some further notation:

$y$ = observed annual sales of innovative products;

| | | |
|---|---|---|
| $y^*$ | = | latent annual sales of innovative products; |
| $c^*$ | = | latent fixed threshold; |
| $x$ | = | exogenous variables that influence the sales of innovative products; |
| $z$ | = | exogenous variables that influence the threshold of sales of innovative products; |
| $\beta$ | = | vector of regression coefficients of variables *x*; |
| $\alpha$ | = | vector of regression coefficients of variables *z*; |
| $\varepsilon$ | = | error term of sales equation; |
| $\eta$ | = | error term of threshold equation; |
| $\sigma_\varepsilon$ | = | standard deviation of $\varepsilon$; |
| $\sigma_\eta$ | = | standard deviation of $\eta$; |
| $\rho$ | = | correlation coefficient of $\varepsilon$ and $\eta$. |

We do not attempt to estimate the structural model in (3) and (4) directly. From (4) we see that the threshold is

(5) $$\mathrm{c}^* = \frac{\mathrm{c.r}}{\gamma.\left(1-\mathrm{e}^{-\mathrm{r.T}}\right)} < (\mu(\mathrm{s}^{\mathrm{opt}}) + \nu)$$

We express $\mathrm{c}^*$ in (5) and $\mu(s^{opt})$ in (3) as a function of exogenous variables *z* and *x*, respectively. In the sequel, we use the exclusion restrictions of the economic model. In other words, variables that affect the threshold through the fixed cost *c* do not enter the sales equation (3). Hence we obtain the latent regression equations

(6) $$c^* = z\alpha + \eta$$

(7) $$y^* = x\beta + \varepsilon$$

(8) $$\begin{pmatrix}\varepsilon\\ \eta\end{pmatrix} \sim N\left(\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}\sigma_\varepsilon^2 & ..\\ \rho\sigma_\varepsilon\sigma_\eta & \sigma_\eta^2\end{pmatrix}\right)$$

Firms will have sales of new products if and only if the (expected) sales of the new product exceed the threshold $\mathrm{c}^*$. Hence, the latent and observed variables are related by

(9) $$I = I\left(y^* > c^*\right)$$

(10) $$y = \begin{bmatrix} 0 \;\; if \;\; I = 0 \\ y^* \;\; if \;\; I = 1 \end{bmatrix}$$

The probability of positive sales of a new product is

(11) $$Pr(\,I = 1) = Pr\left(y^* > c^*\right) = \Phi\left(\frac{x\,\beta - z\,\alpha}{\sqrt{\sigma_{\eta^*}^2}}\right)$$

where $\sigma^2_{\eta^*} = \sigma^2_\eta + \sigma^2_\varepsilon - 2\rho\sigma_\varepsilon\sigma_\eta$, and Φ is the cumulative distribution function of the standard normal distribution. The expected sales of the innovative product, conditional on the event that they are positive, are

$$(12)\qquad E(y|I=1) = x\beta + \frac{\sigma^2_\varepsilon}{\sigma^2_{\eta^*}} \frac{\varphi\left(\frac{x\beta - z\alpha}{\sqrt{\sigma^2_{\eta^*}}}\right)}{\Phi\left(\frac{x\beta - z\alpha}{\sqrt{\sigma^2_{\eta^*}}}\right)}$$

where ϕ is the density function of the standard normal distribution.

The likelihood function of the threshold model is derived from the joint distribution of the observed variables (*y*, I)

$$(13)\qquad f_{y,I}(I=0) = Pr\left(y^* < c^*\right) = 1 - \Phi\left(\frac{x\beta - z\alpha}{\sigma_{\eta^*}}\right)$$

$$(14)\qquad f_{y,I}(y, I=1) = Pr\left(c^* < y^*, y^* = y\right) = Pr\left(c^* < y^* \middle| y^* = y\right).Pr\left(y^* = y\right) =$$

$$\Phi\left(\frac{(y - z\alpha) - \rho.\sigma_\eta\left(\frac{y - x\beta}{\sigma_\varepsilon}\right)}{\sigma_\eta\sqrt{1-\rho^2}}\right).\frac{\varphi\left(\frac{y - x\beta}{\sigma_\varepsilon}\right)}{\sigma_\varepsilon}$$

Hence the loglikelihood is:

$$(15)\qquad LogL_i = \sum_{i=1}^{N} \left\{ \begin{array}{l} (1-I_i).Log\left[1 - \Phi\left(\frac{x_i\beta - z_i\alpha}{\sigma_{\eta^*}}\right)\right] + I_i.Log\left[\frac{\varphi\left(\frac{y_i - x_i\beta}{\sigma_\varepsilon}\right)}{\sigma_\varepsilon}\right] \\ \\ + I_i.Log\left[\Phi\left(\frac{(y_i - z_i\alpha) - \rho.\sigma_\eta\left(\frac{y_i - x_i\beta}{\sigma_\varepsilon}\right)}{\sigma_\eta\sqrt{1-\rho^2}}\right)\right] \end{array} \right\}$$

It is also possible to estimate the model by a two-stage estimation method (see Maddala (1983, pp. 228-230)). We used the estimates of the two-stage method as starting values for the maximum likelihood estimation.

Nelson (1977, p. 315) and Maddala (1983 p. 229) discuss the identification of the parameters of the threshold model. A sufficient condition for identification is that the sales equation contains at least one exogenous variable that is not in the threshold equation. As noted before, variables that affect the threshold through the cost of introduction can be excluded from the sales equation, so that this condition can be easily satisfied.

# 3 The data

## 3.1 Source and background

We use data from the Dutch part of the *Community Innovation Survey* (CIS) for the years 1994-1996 (CIS2) and 1996-1998 (CIS2½, see CBS 1998, 2000). We merged the two datasets in order to have information on a longer period of time and to be able to construct lagged exogenous variables, which will be discussed later on. Both questionnaires consisted of two parts. In the first part, firms were asked to report basic information on the firm, such as the branch of principal activity, sales, exports, employment, etc. The second part contained questions on R&D, innovation and related issues. Only firms that answered at least one of the following three questions in the affirmative were asked to complete the second part of the questionnaire:

- Did your firm develop any technologically changed products during 1996-98?
- Did your firm develop any technologically changed processes during 1996-98?
- Does your firm plan to develop any technologically changed products or processes in the years 1999-2001? Firms affirming one or more of these questions in both questionnaires are defined as being innovative.

In the survey, a distinction is made between three types of product (1) essentially unchanged products, (2) incrementally improved products and (3) radically changed or totally new products. In this paper we will refer to the last category as 'innovative products' and is our object of analysis. In the survey a further distinction is made between:

- Products 'new to the firm';
- Products 'new to the market'.

Although it is tempting to use products 'new to the market' as an indicator of the most innovative products we will use products 'new to the firm' as the endogenous variable because the indicator 'new to the market' suffers from a serious problem in the way firms are interpreting their scope of the relevant market. For instance are firms referring to the Dutch market, the European or the global market? Brouwer & Kleinknecht (1996) point to the relationship between the export quote of firms and the turnover out of products ‘new to the market’, the higher the export quote the less they report of having turnover out of products ‘new to the market’. From this it seems that firms more orientated to the home market, which tends to be relatively small in the Dutch case, are overestimating their innovative efforts. To avoid this ambiguity we restrict our analysis to sales new to the firm.

### 3.2 Descriptive statistics

From the merged dataset firms in the service sector were excluded first because of the fact that those firms do not produce products but services, which are outside the scope of this paper. Next firms were selected meeting the following criteria: presented in both questionnaires (CIS2 and CIS2½), being innovative in 1994-1998 and have R&D output in at least 1996-1998, being industrial firms with 10 employees or more. Finally this selection resulted in a database of 2279 firms. A brief comparison of firms selected and those excluded from the sample show that small firms are underrepresented in the sample, and the selected firms on average have a higher exportquote (see the appendix). Within this sample of 2279 firms we distinguish two groups of firms. The first group comprise of 1002 innovative firms with turnover out of innovative products in 1996-1998 as defined earlier. The second group are innovative firms without turnover out of innovative products in 1996-1998. The latter group comprise of 1277 firms focussing on process innovations or firms in the stage of developing new products not yet introduced onto the market.

The descriptive statistics for all innovative firms are given in Table 1. Table 1 shows that the average turnover is approximately 31 million Dutch Guilders or 16 million US Doller.[3] The difference in turnover between firms with and without turnover out of innovative products is small and insignificant. Also the number of employees does not differ significantly between the two groups of firms. The average size is about a hundred employees. The standard deviations of total turnover and the number of employees are quite large because about 90 percent of all Dutch firms have 50 employees or less but also that there are a few large (multinational) companies in the sample which have high labour productivity and high sales of new products. Thus this reflects the very skewed distribution of size of firms in the Netherlands. Because of this and for theoretical considerations on which we will elaborate later on we will model size in a non-linear fashion. On average innovative firms obtain about 13% of their turnover out of the sale of

[3] In 2000 the Dutch Guilder was replaced by the Euro. All the statistics before 2000 are in Dutch currency. In the year 1998 the average exchange rate was 1.9825 Guilder for one US Dollar (source Dutch National Bank, DNB).

innovative products and approximately 8.3% if weighted with the size of firms meaning that larger firms have less turnover out of new products.

Table 1: Descriptive statistics of the variables used in the model

| | (1) All innovative firms | | (2) Firms with turnover new products | | (3) Firms without turnover new products | | |
|---|---|---|---|---|---|---|---|
| | mean | std.dev. | mean | std.dev. | mean | std.dev. | Sig* |
| Dummy: 1 = turnover out of products new to the firm 1998 | 0.440 | 0.496 | 1 | | 0 | | |
| Log turnover 1998 | 17.269 | 1.464 | 17.241 | 1.491 | 17.292 | 1.443 | |
| Log turnover out of products new to the firm 1998 | 6.487 | 7.423 | 14.754 | 1.818 | n/a | n/a | |
| Turnover out of products new to the firm as % of total turnover 1998 | 0.059 | 0.116 | 0.133 | 0.144 | n/a | n/a | |
| Log number of employees (FTE) 1998 | 4.580 | 1.094 | 4.604 | 1.134 | 4.562 | 1.062 | |
| R&D employment intensity 1996 | 0.024 | 0.056 | 0.041 | 0.067 | 0.014 | 0.047 | *** |
| % of total R&D spending which have to be recovered 1996 | 0.385 | 0.392 | 0.394 | 0.366 | 0.378 | 0.411 | *** |
| Source of information: internal (score 0..1) 1996 | 0.396 | 0.310 | 0.536 | 0.241 | 0.338 | 0.317 | *** |
| Source of information: suppliers or customers (score 0..1) 1996 | 0.324 | 0.284 | 0.449 | 0.241 | 0.268 | 0.283 | *** |
| Source of information: publicly available information, proxy for knowledge spillovers (score 0..1) 1996 | 0.193 | 0.190 | 0.279 | 0.180 | 0.152 | 0.177 | *** |
| Dummy: 1 = firms established in 1994-1996 | 0.022 | 0.148 | 0.024 | 0.153 | 0.021 | 0.144 | |
| Dummy: 1 = firms engaged in R&D-cooperation 1996 | 0.270 | 0.444 | 0.392 | 0.489 | 0.211 | 0.408 | *** |
| Dummy: 1 = firms received subsidies to stimulate technological progress 1998 | 0.428 | 0.495 | 0.626 | 0.484 | 0.273 | 0.445 | *** |
| Observations | 2279 | | 1002 | | 1277 | | |

*Difference between means of firms in dataset 2 and dataset 3: *** = significant at the 1% level, ** = significant at the 5% level and * = significant at the10% level*

The descriptive statistics of the variables for dataset 1 (all innovative firms), 2 (firms with turnover out of new products) and 3 (firms without new products) are different, except for the following variables: turnover, the number of employees and the fraction of firms established in 1994-1996. Dataset 3 has the lowest values and dataset 2 the highest values for these variables.

## 4 Hypotheses

Before we present the results we formulate the hypotheses to be tested with our model. This allows us to identify variables that can be included or excluded from either the fixed cost equation or sales equation. First we deal with hypotheses regarding the sales equation, after that we deal with the fixed cost equation.

**Determinants of sales of new products**

In general the market of new products like any other market is determined predominantly by supply, demand and market conditions, next to other determinants like institutional factors or progress of technology, which we leave aside. In principle firms are only capable to adapt to changes in market conditions by changing their own products i.e. the firm specific supply characteristics and like the entrepreneurs we assume demand factors as given.

The same could be argued about market conditions, only very large firms are in the position to use their market power to mould the market conditions to their interests; the majority of smaller firms have to accept market conditions as given. In the case of market conditions however the situation is more complicated in how decision-making firms deal with competition and which indicators will influence the sales of new products. On the firm level firms compete on product characteristics, price, quality and other firm / product specific factors, but on a more aggregate level competition is the resultant of technological development and demand articulation which can be understood using the theoretical ideas of the product life cycle model which brings together the combined effects of technology development, dominant design, demand, entry and exit, the distribution of size of firms, and market conditions in time.[4] For instance if in a specific

---

[4] The model of the product life cycle builds on the work of Gort & Klepper (1982) and Klepper & Graddy (1990). See among others Jovanovic & MacDonald (1994) and Klepper & Simons (2000), Klepper (2002) for an empirical assessment. The notion of dominant design stems from Albernathy and Utterback (1978), Rosenberg (1982) and Sahal (1986). For a review of the notion of dominant design in economics see Nelson (1995). A different strand of literature points at the relationship between business cycles and innovation using the 'demand pull' and 'supply push' paradigm, and the associated controversy (see Schmookler, 1996, Scherer, 1982). The causal effects of demand pull or supply push on innovative activities are not substantiated by empirical findings however (Scherer, 1982, Geroski & Walters, 1995).

branch there are many opportunities for technological development and improvements and thus a strong articulated demand, firms are forced to invest more in the development of new products because older products are superseded by superior new product (of competing firms). On average the percentage of turnover out of new products will be higher compared to branches with lower technological opportunities. In branches characterised by mature technology and modest opportunities for further technological development firms are more likely to improve products on an incremental basis, and their earnings will stem predominately from improved or unaltered products (but probably cheaper produced because of cost reducing process innovations).

H1: *New established firms are a potential bias?*

The sample includes mature firms and new established firms. First of all we have to correct for start-up firms because by necessity new or start-up firms only have sales from new products and no sales out of improved or unaltered products. Therefore we include in the sales equation a dummy if a firm is a start-up-firm i.e. a firm established in the 1994-1996 era or not in order to correct for this potential bias. We expect this dummy to have a positive effect on the level of sales.

H2: *R&D-intensity and sales of new products.*

Since R&D-expenditures are a major input into the innovation process, we expect a firm's R&D intensity to be related directly to its 'output' of product innovations. R&D intensity is a proxy for firm specific adaptation to market conditions and branch specific technological change.

H3: *Competition / Market structure and sales of new products*

According to the product life cycle model in the first stages of technological development there are many small firms competing on design and product / market combinations. In this situation the market structure shows a low seller concentration. The prospect of becoming a market leader if a firm succeeds in establishing a dominant design or product-market combination, which will attract many adventurous entrepreneurs. This is called the period of experimentation with various product market combinations. This period ends when a certain product-market combination begins to dominate technological development and consumers are more inclined towards a particular design. Firms are forced more or less to abandon their not so successful product-market combination in favour of a more successful design of a competitor.

Later on when a dominant design has been established firms start competing on price and economies of scale becomes an important determinant to survive, starting the 'oligopolistic shake-out'. Many firms who fail to achieve a minimum efficient scale of production sooner or later have to exit, resulting in a market dominated by a few large firms competing on price

(Klepper 1996). Firms facing a Bertrand competition however have a strong incentive to increase their profit margin by means of product differentiation and thus resulting in more sales out of new products if their products succeed over time (Martin, 1993, Kanovski, 2005). The larger incumbent firms will invest heavily in process innovations maximizing the effects of economies of scale in order to sustain their dominant position and to lower the incentives for rival firms to enter the market.

Therefore we expect an (stylized) U-shape relationship between sales of new products and the seller concentration expressed as Herfindahl-equivalent index.[5] This U-shape is also congruent with the Schumpeter mark I and II relationship between the size of the firm and the incentive to innovate (Baumol, 2002, Kamien & Schwartz, 1975, 1982, Scherer, 1962, 1992).[6]

**Determinants of the fixed cost equation**

H4: *Product related R&D investments and the fixed cost equation.*

In the advent of the introduction of a new product onto the market we expect that firms shift their total R&D expenditures towards more product-related R&D-expenditures such as expenditures on development of production facilities or specialised equipment, design and marketing and away from expenditures on fundamental or basic research according to the chain-linked model of innovation of Kline & Rosenberg (1986). The exogenous variable Rdpintr96, defined as the percentage of total R&D-investment spend on the introduction of new products captures this shift in R&D expenditures and we expect that the more firms have to spend on product-development related R&D-activities the higher the costs of introduction will be, that means the costs that have to be recurred. These costs differ from R&D-expenditures in the pre-competitive phase of product-development, which can be considered as sunk costs. Sunk costs by definition are irrelevant in the decision making process whether to introduce new products onto the market or not.

H5: *R&D co-operation and R&D subsidies lower the Fixed Costs of sales of new products*

The sharing of knowledge and risks through R&D co-operation agreements decreases the fixed cost of introduction of new products, besides enhancing future innovation benefits by internalising positive external effects (de Bondt, 1996). R&D subsidies lower the fixed costs directly.

5 The Herfindahl-equivalent is equal to the reciprocal of the Herfindahl index (1/H). The Herfindahl index is the sum of the squared market shares of firms in a particular sector.

6 Recently Aghion et al (2005) proposed a different theoretical model to explain an inverted U-shape between the innovation activities and competition. Their model will be discussed in the result section.

H6: *Knowledge spillovers and the fixed cost equation*

Innovation is all about developing (technical) knowledge. In his seminal paper Jaffe (1986) states that: 'Since knowledge is inherently a public good, the existence of technologically related research efforts of other firms may allow a given firm to achieve results with less research effort than otherwise', pp. 984.[7] From this we expect that using external sources of knowledge such as information from consultants, universities, patent-data, and other semi-public agencies (brinf_96) will lower the costs of R&D or more in particular the fixed costs of introducing new products onto the market. Besides knowledge spillovers originating from (semi) public sources, we distinguish two other sources of knowledge. The first one is internal sources of information i.e. information rendered inside the firm (brint_96). The second one is sources of information originating from suppliers and customers (brrel_96). According to von Hippel (1988), it is not necessarily the inventor who will develop the innovation. Instead the actor in the vertical chain of product development (distributed innovation process) who has the best opportunity to manage the appropriability conditions will be the innovator. We expect that firms who rely on information of suppliers and customers as sources of innovation have a better opportunity to appropriate the rents of innovation and will have to spend less on costs on introduction of new products (see among others Lilien et.al., 2002). For instance if the source of innovation is from a lead customer/user, the innovating firms will face less market uncertainties and thus have to spend less in marketing costs compared to a firm who relies predominately on internal generated ideas.

**Other exogenous variables**

H7: *Size of firms and innovative products*

An important exogenous variable is the size of the firms, operationalized as the log of the number of employees. We expect the size of the firm will affect the level of sales and the amount of investments to introduce new products onto the market both. The larger the firms the more resources are available to invest in the introduction of new products and at the same time larger firms are expected to have better market opportunities to realise sales out of new products. This combined effect resembles the theory of 'spreading' introduced by Cohen & Klepper (1996) to explain why larger firms invest relatively more in R&D albeit with a lower efficiency compared

---

7 See also Bernstein & Nadiri (1988) who distinguish between receiving and sending industries within a spillover network. From the point of view of individual firms, a firm tends to focus on receiving knowledge spillovers and at the same time try to avoid sending spillovers and thus creating positive appropriability conditions (Teece, 1986). CIS2 does not comprise direct information to distinguish between incoming (receiving) and outgoing (sending) knowledge spillovers, especially the measurement of outgoing knowledge spillovers is cumbersome (see Cassiman & Veugelers, 2002, in a attempt to use proxies to distinguish between incoming and outgoing knowledge spillovers).

to smaller firms. The explanation of Cohen & Klepper boils down to the fact that due to their market share and distributions and marketing efforts larger firms need less successful innovative products to cover their initial R&D-investments and for that matter small firms with very limited resources can not bear the risk of market failure of new products and therefore have to choose their R&D-investment more carefully, resulting in more efficient R&D-investments, i.e. a higher R&D-investment new product ratio, conditional on succeeding to tackle the hurdle of new product development. The situation is a bit more complicated though. On average larger firms will have a larger product portfolio and to prevent 'cannibalizing' their existing product portfolio larger firms will be more reluctant to introduce new products onto the market (Conner, 1988). We will expect that size will have a non-linear effect on both the level of sales and on the fixed cost. Initially when firms are relatively small, the size effect will have a strong positive effect, but when firms become larger the positive effect of being large will diminish and due to the 'cannibalizing' effect on the existing product portfolio.
Further on we include branch dummies in the cost equation in order to correct for technology and or branch specific cost of introduction.

# 5 Results

The estimates of the threshold model are summarised in Table 2 for products new to the firm. The threshold is not observed directly but can be derived from the model by subtracting the costs of introduction from the sales. A positive coefficient in the costs equation means that the corresponding variable is positively related to the sales threshold. In other words, it increases the threshold. If the coefficient of a variable that appears in the costs equation *and* in the sales equation is positive and if it is larger in the sales equation than in the threshold equation, then the variable increases the probability that the firm innovates (see equation 11). The same is true if the coefficient in the sales equation is negative and smaller than in the threshold equation.

Even though the threshold is not observed, the parameters of the threshold equation and those of the sales equation are identified by exclusion restrictions. The theoretical model of section 2.1 implies that variables that affect the fixed costs can be excluded from the sales equation. For that reason a variable such as R&D collaboration (partne96) or product introduction related R&D-expenditures (rdpintr96) do not enter the sales equation. We do allow for correlation between the error terms in the threshold and sales equations. This correlation captures common firm specific variables that have been omitted from both equations.

Table 2: The estimation results

| | Exogenous variables | Sales equation coefficient | t-value | | Fixed Costs equation coefficient | t-value | |
|---|---|---|---|---|---|---|---|
| constant | constant | 11.004 | 185.020 | *** | 10.699 | 170.853 | *** |
| dstart96 | Dummy: 1 = firms established in 1994-1996 | 0.337 | 5.524 | *** | | | |
| rdi_96 | R&D employment intensity 1996 | 1.670 | 25.652 | *** | | | |
| lnw_98 | Log turnover 1998 | 0.813 | 28.214 | *** | 1.468 | 28.214 | *** |
| lnw298 | Log turnover 1998, squared | 0.303 | 7.513 | *** | -0.231 | 7.513 | *** |
| he3d98 | Herfindahl equivalent index | -0.195 | -6.611 | *** | | | |
| he3d982 | Herfindahl equivalent index squared | 0.237 | 7.414 | *** | | | |
| rdpintr6 | % of total R&D spending which have to be recovered 1996 | | | | 0.134 | 2.360 | ** |
| partne96 | Dummy: 1 = firms engaged in co-operation 1996 | | | | -0.036 | -0.649 | |
| subsid98 | Dummy: 1 = firms received subsidies to stimulate technological progress 1998 | | | | -1.017 | -17.997 | *** |
| brint_96 | Source of information: internal (score 0..1) 1996 | | | | -0.466 | -7.621 | *** |
| brrel_96 | Source of information: suppliers or customers (score 0..1) 1996 | | | | -0.475 | -7.740 | *** |
| brinf_96 | Source of information: publicly available information, proxy for knowledge spillovers (score 0..1) 1996 | | | | -1.117 | -17.444 | *** |
| sbi_pdj2 | Dummy: 1 = Pavitt taxonomy, Scale-intensive industrial firms # | | | | -1.244 | -21.515 | *** |
| sbi_pdj3 | Dummy: 1 = Pavitt taxonomy, Specialised equipment suppliers # | | | | -1.139 | -18.844 | *** |
| | Nr. obs. | 2279 | | | | | |
| | Loglikelihood | -2892.82 | | | | | |
| | Loglikelihood base-line model | -3583.59 | | | | | |
| | LR test | 1381.54 | | | | | |
| | Mc Fadden $R^2$ | 0.19 | | | | | |
| | Zimmerman-Veal $R^2$ | 0.50 | | | | | |

**** = Significant at the 1% level, ** = significant at the 5% level and * = significant at the10% level*

[#] Adopted from Pavitt (1984).

Next, we discuss the parameter estimates in tables 2 starting with the effect of size on both the fixed costs and sales, followed by the parameter estimates of the sales equation and last the parameter estimates of the fixed costs equation.

*Size*

As expected lager firms realise higher sales of new products. Larger firms benefit from economies of scale and scope and thus have the production and marketing capacity to realize higher sales of new products. Smaller firms are forced to focus on niche markers, which by definition are smaller, and thus the potential to realise sales of new products is smaller as well.

**Parameter estimates of the sales equation**

Table 2 shows that new firms (dstart96) have indeed a significant higher percentage of sales of new products thus supporting hypothesis 1 stating that the inclusion of new established firms could introduce a bias in the estimations.

A higher R&D intensity (rdi_96) results in higher sales of new products. This is consistent with the economic model of section 2 and hypothesis 2 above.

The effect of market structure measured by a Herfindahl-equivalent index (he3d98) on the sale of new products, departs from hypothesis 3. Instead of a U-shape, we find a non-linear negative relationship (a hockey stick on its back, with the tip of the stick facing upwards) between the seller concentration and sales of new products meaning that markets with a lower seller concentration i.e. dominated by relatively small firms generate more turnovers out of new products. Our empirical result is in congruence with most of empirical work on the relationship between competition and innovation (see Nickell, 1996), and this finding collaborates with the Schumpeter mark I theory (Schumpeter, 1934) and the Arrow's replacement effect, which restrain large companies to invest heavily in product innovations (Arrow, 1962). However this result does not comply completely with the proposed Schumpeterian and Product-life cycle models, put forward in the hypothesis section of this paper, which predict a U-shape relationship between innovative activities and market structure. Recently the relationship between competition and innovation became more puzzling. Agion et al. (2005) show that an inverted U-shape relationship between competition and innovation fit UK panel data covering the period of 1973-1994 much better, challenging both empirical and theoretical findings of the traditional Industrial Organization (IO) literature, and contraire to our theoretical and empirical findings. In their paper they present an alternative model which boils down to different strategic step-by-step product innovation behaviour in so called levelled and neck-and-neck industries. The 'escapecompetition effect' increases the incremental profit from innovation, but the 'Schumpeterian effect' reduces innovation incentives for laggards as competition increase. "The balance between these two effects changes between low and high levels of competition, generating an inverted-U relationship", (Aghion et al, pp. 720-721).

**Parameter estimates of the fixed costs equation**

*Product related R&D-investments*

Hypothesis 4 stated that there is a positive relationship between product-related R&D-expenditures as a percentage of total R&D-investments and the fixed costs, which have to be recurred from sales of new products. The parameter estimate confirms this hypothesis showing a positive and significant coefficient. We also stated that according to economic theory sunk costs do not have any significance when to decide on investments like R&D-expenditures on the introduction of new products onto the market.

*R&D collaboration and R&D-subsidies*

In correspondence with hypothesis 5 R&D collaboration (partne96) and subsidies (subsid98) reduce the fixed costs of introduction although the effect of R&D-collaboration is insignificant. In the Netherlands a major R&D policy measure is a general tax-credit facility to lower the labour costs of R&D-personnel called the WBSO.[8] The purpose of the WBSO is to reduce the R&D costs of firms, in particular those of SME's. Large firms are bound to a maximum of tax-credits. Subsidies including the WBSO are decreasing the costs of introducing new products onto the market. One could argue that the effect of the WBSO may not entirely exogenous. Firms anticipating the introduction of new products may have an incentive to apply.

A possible explanation for not finding a significant effect of collaboration on the fixed costs could be that besides a costs reducing motive to be engaged in R&D-collaboration, there could be a skill sharing motive as well such as to develop and share new knowledge different fields of technology or different markets (Sakakibara, 1997). The estimates suggest that the cost-reducing motive is less important. That does not point automatically to the conclusion that the skill-sharing motive is dominant. A skill-sharing motive may lead to the reduction of R&D-costs and thus the effect of a skill-sharing type of R&D-collaboration would not differ from a costs reducing type of R&D-collaboration regarding the fixed costs. Perhaps a better explanation could be to look at the type of knowledge the R&D-collaboration is focused on. The collaborative efforts could be focused on applied knowledge for specific products competing on the market (non-collusive R&D-cooperation) or geared towards more basic knowledge to strengthen the innovative skills of the collaborators on a basic level and in an early phase of the innovation cycle. In the latter case the effect of R&D-cooperation will have a (much) smaller or even insignificant effect on the fixed costs. We do not have additional information, which explanation is more plausible, but it seems that the skill-sharing type of R&D-collaboration is predominate.

*Knowledge spillovers*

Table 2 shows that all sources of information (brint_96, brrel_96 and brinf_96) lower the fixed costs significantly. The estimated coefficients show that knowledge spillovers (brinf_96) have the strongest and the use of internal knowledge the weakest effect. In concordance with Jaffe (1986) this result shows that innovating firms benefit from overall knowledge spillovers most. The question is how to interpret these findings. Jaffe (1986) argues that knowledge spillovers reflect the technological opportunities that exist in a particular branch. Technological opportunities are an incentive for product development and a signal that product differentiation is a profitable strategy. If a particular branch is low on technological opportunities the amount of research needed to come up with something new will be higher compared to a branch with an abundance of technological opportunities. According to Jaffe's reasoning it is not so much the knowledge itself as well as its signalling effect. That would explain why the coefficients of other sources of knowledge are in the same magnitude albeit on a lower level compared to knowledge spillovers (brinf_96). The estimates of brint_96, brrel_96 reveal something about the innovating strategies: firms rely on their own technological strength (technology initiators) or are more involved into developments in a distributed product development exposed by von Hippel (technology followers). Both strategies appear to be equally effective.

Finally we have to mention the fit of out model is rather well. The Zimmerman Veal R2 and the Mc Fadden R2 are 0.50 and 0.19 respectively. The value of the logarithm of likelihood of the model and the baseline model are respectively –2892 and –3583.

## 6 Summary and conclusions

In this paper, we propose an economic model to analyse the sales out of new products. This model accounts for the fact that even among firms for which R&D is a permanent activity, a fraction of firms does not have sales of innovative products during a two-year observation period. In this study we analyse an indicator of a firm's innovative output 'new to the firm'. In this study we start from the assumption that, in order to introduce a product innovation, the firm

---

[8] WBSO is an acronym of the Dutch name of the act to encourage R&D (Wet Bevordering Speur- en

must incur fixed costs. At the time of considering a new product introduction, the firm will compare fixed costs with expected (net) revenues during the economic lifetime of the product, and only if revenues exceed fixed costs (threshold), the firm will introduce the product.

Our results show that firms face a threshold when deciding on whether or not to introduce new products onto the market. The fixed costs of introduction impose a hurdle to introduce new products. Firms are focused on means to lower those costs. Among the factors, which are of importance to lower the fixed costs are knowledge spillovers and R&D related subsidies. It comes at no surprise that R&D-related subsidies decrease the fixed costs. Of more interest are the differences between knowledge spillovers and sources of innovation and their effect on the fixed costs. Those differences reveal much about the knowledge management strategies of innovating firms. In general spillovers of public knowledge have a profound effect on the fixed costs, but interestingly knowledge spillovers originating from the firm itself or from suppliers and customers also have an equally decreasing effect on the fixed costs albeit at a lower level. This point at the process of the distributed innovation process described by von Hippel (1988) stating that the inventor and innovating actor do not necessarily coincide but who will be the innovator depends on which actor along the value chain is able to appropriate the economic rents of an innovation. The results show that both strategies of technology initiator and technology follower appear to be equally effective.

We also find conformation of the idea that in the period preceding the introduction firms shift their R&D-expenditures towards market related activities such as expenditures on development of production facilities or specialised equipment, design and marketing and away from expenditures on fundamental or basic research according to the chain-linked model of Kline & Rosenberg (1986). The level of sales of new products depends on the R&D-intensity and market structure ones more substantiating the Schumpeter mark I paradigm and the Arrow replacement effect, meaning that small (high-tech) firms are more important to commercialize innovative products, partially substantiating the product-life cycle model of Cohen & Klepper (1996).

The empirical results of Aghion et al (2005) and the proposed theoretical model to explain the inverted-U shape relationship between innovation and competition, although dealing with a slightly different topic we want to address in this paper, ones more show that the understanding of the relationship between incentives to innovate and competition leaves much to be desired.

---

Ontwikkelingswerk)

# 7 Literature

Aghion, P., N. Bloom, R. Blundell, R. Griffith & P. Howitt (2005), 'Competition and Innovation: An Inverted U Relationship', The Quarterly Journal of Economics (120-2), pp. 701-728

Albernathy, J.M. and Utterback, J.M. (1978), 'Patterns of Industrial Innovation, Technology Review (41-7), pp. 2-9

Arrow, K.J. (1962), 'Economic welfare and the allocation of resources for invention'. In: R.R. Nelson (ed), The Rate and Direction of Inventive Activity: Economic and Social Factors, Princeton University Press for the NBER

Baumol, W.J. (2002), The free market innovation machine: analysing the growth miracle of capitalism. Princeton, Princeton University Press

Brouwer, E. (1997), 'Into innovation: determinants and indicators', Ph.D. Dissertation, University of Amsterdam, The Netherlands.

Brouwer, E. & A.H. Kleinknecht (1996), 'Determinants of innovation: a micro econometric analysis of three alternative innovation output indicators'. In: Determinants of innovation. The message from new indicators, A. Kleinknecht, (ed.), London: Macmillan Press, pp. 99-124.

Cassiman, B. & R. Veugelers (2002), 'R&D Cooperation and Spillovers: Some Empirical Evidence from Belgium', American Economic Review (92), pp. 1169-1184

CBS (1998), Kennis en Economie 1998, Statistics Netherlands, Voorburg/Heerlen, CBS

CBS (2000), Kennis en Economie 2000, Statistics Netherlands, Amsterdam, Elsevier bedrijfsinformatie

Church, J. & R. Ware (2000), Industrial Organization, A strategic Approach, Boston , Irwin McGraw-Hill

Cohen, W.M. & S. Klepper (1996), 'A Reprise of size and R&D', The Economic Journal (106-437), pp.925-951

Conner, K.R. (1988), 'Strategies for product cannibalism', Strategic Management Journal (9), pp. 9-26

De Bondt, R. (1996), 'Spillovers and innovative activities', International Journal of Industrial Organization (15), pp. 1-28

Geroski, P.A. & C.F. Walters (1995), 'Innovative activity over the business cycle', Economic Journal (105), pp. 916-928.

Gort, M & S. Klepper (1982), 'Time path in the Diffusion of Product Innovations' Economic, Journal (92), pp. 630-653

Hippel, E. von (1988), The sources of innovation. New York/Oxford, Oxford University

Jaffe, A.B. (1986), 'Technological Opportunity and spillovers of R&D: Evidence from firms' patents, profits, and market value', The American Economic Review, (76-5), pp. 984-1001

Jovanovic, B. & G.M. MacDonald (1994), 'The Life Cycle of a Competitive Industry', The Journal of Political Economy (102-2), pp. 322-347

Kamien, M.I. & N.L. Schwartz (1975), 'Market Structure and Innovation: A Survey', Journal of Economic Literature (13-1), pp. 1-37

Kamien, M.I. & N.L. Schwartz (1982), 'Market structure and innovation', Cambridge, Cambridge University Press.

Kaniovski, S. (2005), 'Product differentiation and competitive selection', Journal of Evolutionary Economics (15), pp. 567-580

Kleinknecht A.H. (ed.) (1996): Determinants of innovation, the message from new indicators, London, Macmillan Press.

Klepper, S. (1996), 'Entry, Exit, Growth, and Innovation over the Product Life Cycle', The American Economic Review (86-3), pp. 562-583

Klepper, S. (2002), 'Firm Survival and the Evolution of Oligopoly', The RAND Journal of Economics (33-1), pp. 37-61

Klepper, S. & E. Graddy (1990), 'The Evolution of New Industries and the Determinants of Market Structure', Rand Journal of Economics (21), pp. 24-44

Klepper, S. & K.L. Simons, (2000), 'The Making of an Oligopoly: Firm Survival and Technological Change in the Evolution of the U.S. Tire Industry', The Journal of Political Economy (108-4), pp. 728-760

Kline, S.J. & N. Rosenberg (1986), 'An overview of innovation'. In: R. Landau & N. Rosenberg (eds), The positive sum strategy: harnessing technology for economic growth, Washington, National Academy Press, pp. 275-307

Lilien, G.L., P.D. Morrison, K. Searls, M. Sonnack & E. von Hippel (2002), 'Performance Assessment of the Lead User Idea-Generation Process for New Product Development', Management Science (48-8), pp. 1042-1059

Maddala, G.S. (1983), Limited-dependent and qualitative variables in econometrics, Cambridge, Cambridge University Press.

Martin, S. (1993), Advanced industrial economics, London, Blackwell

Nelson F.D. (1977), 'Censored regression models with unobserved, stochastic censoring thresholds', Journal of Econometrics (6), pp. 309-327.

Nelson, R.R. (1995), 'Recent Evolutionary Theorizing About economic Change', Journal of Economic Literature (33-1), pp. 48-90

Nickell, S.J. (1996), 'Competition and Corporate Performance', Journal of Political Economy (104-4), pp. 724-746

Pavitt, K. (1984): 'Sectoral patterns of technological change: Towards a taxonomy and a theory', Research Policy (13), pp. 343-373

Poot, T. (2004), Determinants of knowledge intensive R&D-collaboration, Ph.D. Dissertation (in Dutch), Delft University of Technology, The Netherlands.

Rosenberg, N. (1982), Inside the Black Box: Technology and Economics, Cambridge, Cambridge University Press

Sahal, D. (1986), Patterns of technological innovation. Reading Mass., Addison-Wesley

Sakakibara, M. (1997), 'Heterogeneity of firm capabilities and cooperative research and development: an empirical examination of motives', Strategic Management Journal (18, summer special issue), pp. 143-164

Scherer, F.M. (1967), 'Market Structure and the Employment of Scientists and Engineers', American Economic Review  (57), pp. 524-531

Scherer, F.M. (1982), 'Demand-pull and technological invention: Schmookler revisited', Journal of Industrial Economics (30), pp. 225-237.

Scherer, F.M. (1992), 'Schumpeter and Plausible Capitalism', Journal of Economic Literature (30), pp. 1416-1433

Schmookler, J. (1966), Invention and economic growth, Cambridge, Harvard University Press.

Schumpeter, J.A. (1934), The theory of Economic Development, Cambridge (Mass), Harvard Economic Studies

Teece, D.J. (1986), 'Profiting from technological innovation: Implications for integration, collaboration, licensing and public policy', Research Policy (15), pp. 285-305

## Appendix, Data preparation and selection

Economic theoretical considerations let to the decision to incorporate lagged exogenous variable in the model. To create a dataset with lagged variables we make use of two questionnaires. The first one is the regular Dutch CIS questionnaire comparable with the Eurostat version, which comprise data of the periode 1994-1996. Statistics Netherlands has the policy to repeat the CIS questionnaire after 2 years. This 'in between' questionnaire, covering the period of 1996-1998, is identical with the Eurostat version, although extended with some additional questions, which are of no concern for the purpose of this paper.

The initial dataset has been constructed by merging the Dutch CIS2 (1994-1996) and CIS2½ (1996-1998) questionnaires. From this database of 10996 firms, three groups of firms were excluded: firms in the service sector, firms without any innovative activity according to the Frascati manual, and firms who are present in only one period. After the exclusion of those firms we are left with a database, which consists of 2278 firms of which we have information covering the whole period of 1994-1998. Those 2278 firms are the selection used to analyse the innovation threshold. Of those 2287 firms there are 1001 firms with sales out of new products and 1277 without sales, i.e. firms assumed not be able to meet the threshold requirements.

The main question is if the 2278 firms used in our analysis are representative for all firms. In order to analyse this potential selection bias we compared the mean values of four firms indicators among the various selections presented earlier. The indicators are: (1) the number of employees and (2) the turnover in 1998, representing different aspects of the size of the firms involved. The two other indicators are (3) the turnover per employee representing a kind of 'productivity' measure, and (4) the export quote representing a competition measure. Firms with a higher export quote are expected to face a stronger competition because they are operating on a global market, while firms with a lower export quote are geared towards the domestic market. Of course these 4 indicators are rough indicators, constructed for the sole purpose of identifying a possible selection bias, and not for a serious analysis of firm characteristics. The descriptives of the four indicators and the results of a t-test whether the means are statistically different from each other are presented in the table below.

Table 1 shows that the firms excluded from the main database are on average significantly smaller, both in terns of the number of employees and turnover. The turnover per employee does not differ significantly but the excluded firms on average have a significantly lower export quote. This selection bias can be explained by the sampling procedure Statistic Netherlands is using. Due to the very skewed distribution of firm size, approximately 82% of all firms in the Netherlands have 10 to 50 employees, a random selection of approximately 18% of the firms with 10 to 50 employees receive a CIS questionnaire, while all firms with 50 or more employees receive a questionnaire. Merging two datasets (CIS2 and CIS2½) small firms have a smaller propensity to be selected in both CIS questionnaires (3% ≈ 0.18*0.18), and selecting only those firms, which are present in both questionnaires, causes a selection bias. To put it differently, merging two CIS questionnaires and selecting firms present in both questionnaires will favour the larger firms (i.e. 50 employees and more). This is reflected in the results of table 1, showing that the selected firms, which are present in both questionnaire are on average larger.

Table 1 Descriptives of 4 firm indicators: number of employees, turnover, turnover per employee, and export quote in 1998 after merging CIS2 and CIS2½

| Whole database (CIS2 and CIS2½ merged) | | | Number of employees (fte) 1998 (1) | Turnover 1998 (x 1000 guilders) (2) | Turnover (x 1000 guilders) per employee (1998) (3) | Export quote 1998 (4) |
|---|---|---|---|---|---|---|
| nr. obs. | 10996 | mean | 131.47 | 66138.55 | 521.216 | 11.3% |
| | | std error mean | 7.05 | 3473.37 | 31.973 | 0.002 |
| firms not selected | | | | | | |
| nr. obs. | 8718 | mean | 106.65 | 49263.96 | 518.463 | 8.6% |
| | | std error mean | 7.46 | 2656.46 | 39.378 | 0.002 |
| firms selected | | | | | | |
| nr. obs. | 2278 | mean | 226.44 | 130734.40 | 531.752 | 21.5% |
| | | std error mean | 18.44 | 13249.91 | 33.320 | 0.007 |
| t-test: selected versus not selected firms | | | | | | |
| | | diff | -119.79 | -81470.42 | -13.289 | -12.9% |
| | | std err | 19.89 | 13513.58 | 51.583 | 0.007 |
| | | t-value | -6.023 | -6.029 | -0.258 | -17.938 |
| | | Prob > \| t\| | 0.0000 | 0.0000 | 0.7967 | 0.0000 |

Table 2 Descriptives of 4 firm indicators: number of employees, turnover, turnover per employee, and export quote in 1998 restricted to the population of firms meeting all selection criteria

| selected firms (nr. obs 2278) | | | Number of employees (fte) 1998 (1) | Turnover 1998 (x 1000 guilders) (2) | Turnover (x 1000 guilders) per employee (1998) (3) | Export quote 1998 (4) |
|---|---|---|---|---|---|---|
| firms without turnover new products | | | | | | |
| nr. obs. | 1277 | mean | 227.44 | 119672.90 | 623.736 | 10.6% |
| | | std error mean | 29.870 | 17679.830 | 56.261 | 0.007 |
| firms with turnover new products | | | | | | |
| nr. obs. | 1001 | mean | 225.17 | 144831.80 | 414.405 | 35.3% |
| | | std error mean | 17.580 | 20013.620 | 23.993 | 0.011 |
| t-test: firm with versus without turnover new products | | | | | | |
| | | diff | 2.27 | -25158.90 | 209.332 | -24.7% |
| | | std err | 34.66 | 26704.33 | 61.164 | 0.013 |
| | | t-value | 0.066 | -0.942 | 3.422 | -18.810 |
| | | Prob > \| t\| | 0.9477 | 0.3462 | 0.0006 | 0.0000 |

As a second step we investigated whether there are differences between firms with and without sales out of new products. Table 2 shows the results. Within the population of selected firms size is not a discriminating factor anymore. However turnover per employee and the export quote differs significantly between the two subpopulations albeit in different directions. Firms with turnover out of new products on average have a lower turnover per employee but a higher export quote.