Centre of Excellence
for Science and Innovation Studies

CESIS Electronic Working Paper Series

Paper No. 392

Persistence of various types of innovation analysed and explained

Charlie Karlsson
Sam Tavssoli

January, 2015

The Royal Institute of technology
Centre of Excellence for Science and Innovation Studies (CESIS)
http://www.cesis.se

Persistence of various types of innovation analysed and explained

Charlie Karlsson[1] & Sam Tavassoli[2]

Abstract: This paper analyses the persistency in innovation behaviour of firms. Using five waves of the Community Innovation Survey in Sweden, we have traced the innovative behaviour of firms over a ten-year period, i.e. between 2002 and 2012. We distinguish between four types of innovations: process, product, marketing, and organizational innovations. First, using Transition Probability Matrix, we found evidence of (unconditional) state dependence in all types of innovation, with product innovators having the strongest persistent behaviour. Second, using a dynamic probit model, we found evidence of "true" state dependency among all types of innovations, except marketing innovators. Once again, the strongest persistency was found for product innovators.

Keywords: persistence, innovation, product innovations, process innovations, market innovations, organizational innovations, state dependence, heterogeneity, firms, Community Innovation Survey

JEL-Codes: D22, L20, O31, O32

[1] Jönköping International Business School (JIBS), Blekinge Institute of Technology & Centre of Excellence for Science and Innovation Studies (CESIS). E-mail: charlie.karlsson@jibs.hj.se

[2] Blekinge Institute of Technology. E-mail:sam.tavassoli@bth.se

1. Introduction

The performance of firms even in the same industry is highly skewed and this heterogeneity in performance is to a high extent persistent over time. Innovation[3] can be seen as one major determinant of the performance of firms, which would imply that the observed heterogeneity in performance among firms actually mirrors persistent differences in innovation behaviour among firms (Geroski, Van Reenen & Walters, 1997). This implies that in every industry we should be able to observe firms that innovate persistently, firms that innovate now and then and firms that never innovate. However, it also implies that firms can survive in an industry also if they choose a strategy not to innovate. Obviously, it is interesting to understand what factors that induces firms to choose strategies implying continuous, intermittent or no innovation (Brown & Eisenhardt, 1998).

Innovation is here seen as the purposeful and intended result of the ability of firms to generate new knowledge and their decisions to apply it to new products and product varieties, processes, organizational designs, combinations of inputs and markets (Fagerberg, Mowery & Nelson, 2005). The persistence of innovation highlights the influence of past and current innovation on future innovation. It has become an important topic in applied industrial economics since the publication of a seminal paper by Geroski, Van Reenan & Walters (1997). The line of empirical research that followed gave rise to an increased conviction that the competitive advantage of firms mainly depends on their ability to innovate over longer periods of time (Le Bas, Mothe & Nguyen-Thi, 2011). However, this ability is a function of environmental, organizational, process and managerial characteristics of firms (Koberg, Detienne & Heppard, 2003). We still have a limited understanding of the long-term determinants of the innovation behaviour of firms including their investments in different types of innovation, such as products, processes, organization and markets. To increase our understanding of these issues, we in this paper try to answer the following four questions: Is innovation persistent at the firm level? Is this true for all types of innovation? If innovation persistence exists, what drives the phenomenon? Are the drivers the same for all types of innovation? Are there complementarities among different types of innovation, i.e. does one type of innovation, e.g. product innovation, induce other types of innovation?

[3] In this paper, we will not discuss the problems of actually defining an innovation since we are using the definitions used in the European Community Innovation Surveys. The definition problem is highlighted in, for example, Garcia & Calantone (2002).

Why are these questions interesting and important? Persistence in innovation has far-reaching effects for various fields of economics dealing with innovation, for the strategic management and operation of innovation processes and for public policy focusing innovation (Peters 2007). Firstly, they are important from the point of view of economic theory. A proven persistence would validate endogenous growth theory, since according to that theory sustainable economic growth is a function of firms' capacity to accumulate economically useful technological knowledge. However, different endogenous growth models make different fundamental assumptions about the determinants of the innovation performance of firms. In the Romer-model, it is assumed that innovation mainly is persistent at the firm level (Romer, 1990). Here it is assumed that incumbent firms and cumulative knowledge creation are the fundamental sources of innovation and economic growth. However, the Romer approach neglects the role of new entrants and creative destruction as drivers of innovation and economic growth and to acknowledge this we have to turn to endogenous growth models including creative destruction processes, which, for example, assume a process of a perpetual renewal of innovators (Aghion & Howitt, 1992). The only way to assess these different representations of the economic growth process and the dynamics in the innovation behaviour of firms is through empirical analyses (Cefis, 2003). Empirical studies are furthermore important for the understanding of the long-term dynamics of industries (Antonelli, Crespi & Scellato, 2012). Secondly, from a strategic management perspective persistence of innovation, i.e. a continuous loop of innovation, supplies a fundamental building block of maintained competitive advantage and long-lived inter-firm performance differences (Ganter & Hecker, 2013). Thirdly, knowledge about the drivers of firms' innovation behaviour is critical for policy makers. If innovation is persistent in the sense that innovation drives innovation, policies designed to support innovation can be expected to have more far-reaching effects since they not only affect innovation in the current period but also in future periods and thus in principle should be able to raise innovation to new levels. Thus, true innovation persistence implies the existence of inter-temporal spillovers, which provides a foundation for the evaluation public programs designed to stimulate innovation. The existence of true innovation persistence also suggests that innovation policies should avoid stimulating the start-up of firms and firms entering new markets. On the other hand, if the observed persistence is the result of other underlying firm characteristics, policy makers should rather try to stimulate those underlying characteristics of firms that drive innovation.

Two mechanisms can explain persistence in innovation of firms. Innovation persistence may be the result of true state dependence and/or spurious state dependence (Heckman, 1981 a & b). True state dependence represents a casual behavioural relationship or if we like a path-dependent process, where the decision to innovate in one period increases the probability to decide and to succeed to innovate in the following period. Such a dependence can be explained by (i) sunk R&D costs (Sutton, 1991), (ii) dynamic increasing returns from knowledge and R&D (Stiglitz, 1987) and/or (iii) earlier success in innovation stimulating further innovation i.e. success bread success (Flaig & Stadler, 1994). The actual probability might of course change over time due to internal and external events and to changing levels of knowledge externalities (Antonelli, Crespi & Scellato, 2013). Spurious state dependence, on the other hand, prevails when the determinants of innovation persistency (e.g. size of firms) are persistent themselves, hence making firms to be more inclined to innovate in a persistent way. Innovation persistence is here the result of the serial correlation in unobservables that generate different innovation competencies and capabilities of firms, i.e. dynamic capabilities (Teece & Pisano, 1994) in line with the resource-based theory of the firm (Penrose, 1959; Langlois & Foss, 1999). However, if these unobservable and serially correlated characteristics (e.g. risk attitudes or managerial skills) are not controlled for in the econometric estimations, they may generate the impression that innovation in one period drives innovation in the following period. Therefore, in reality what is observed is the effect of unobservable characteristics of firms, and not the true persistence of innovation itself.

The introduction above provides a general motivation for more analyses of the persistence of innovation, since the existence of such persistence have strong implications. However, there also exist some more specific motivations to why we should put more effort into the analysis of these phenomena. The first specific motivation is that earlier research in the field has almost exclusively focused on technological innovation and neglected other forms of innovation such as organizational and market innovations. The only major study that we have found that takes a broader perspective is Ganter & Hecker (2013), who use data for Germany. The second specific motivation is that most of the earlier studies say very little about the possible mechanisms underlying persistence of innovation able to discriminate between the two major explanations.

The purpose of this paper is to analyse persistent patterns of innovation for different types of innovation using Swedish data from four waves of Community Innovation Surveys and to test possible explanations for proven persistence. The contribution of this paper is as follows: (i) using the long panel of Community Innovation Survey (CIS) data and tracing the innovative behaviour of firms during ten years period (this is, to our knowledge, the longest panel of CIS that is constructed) (ii) incorporating four types of innovation, i.e. product, process, marketing, and organizational innovations, and (iii) moving beyond the usual manufacturing sector and including the service sector in the analysis as well (Peters (2009) is an exception).

The rest of the paper is organized as follows. Section 2 provides the theoretical causes of innovation persistence. Section 3 offers a short overview on empirical evidence concerning the persistency of innovation. Section 4 shows the data. Section 5 investigates whether there is a persistency in various types of innovation, while Section 6 analyses whether it is a true persistency or not. Section 7 concludes.

2. The Underlying Theoretical Causes of Innovation Persistence

The underlying theoretical causes of innovation persistence are not well understood to put it mildly. However, by consulting a few different fields of economics, we may at least be able to present some different potential causes to why innovation might demonstrate state dependence over time.[4]

2.1 Knowledge, Learning and Dynamic Scale Economies

We start from the economics of knowledge. Already, Geroski, Van Reenen & Walters (1997) suggested that innovation persistence could be explained by a combination of learning effects from the innovation process and positive feed-back mechanisms between the accumulation of knowledge and innovation processes generating dynamic scale economies. Thus, innovation is the result of cumulative knowledge patterns and learning dynamics (Colombelli & von Tunzelmann (2011). Technological knowledge is as an economic good characterized by being cumulative and non-exhaustible (Nelson, 1959; Nelson & Winter, 1981; Ruttan, 1997). At the same time as knowledge is an input in knowledge production process, it is also an output from the same process (David, 1993). These attributes have distinct implications for innovation persistence. The creation of new knowledge vintages have an effect on the disposable knowledge stock that can be used as an input in knowledge generation due to that knowledge is non-

[4] Antonelli (2008) stresses that, it is important to make a distinction between 'path-dependent' and 'past-dependent' innovation persistence. If current innovation can be explained by past innovation, we have 'past-dependent' innovation persistence. If, on the other hand, current innovation is a result of processes determined by initial conditions, we talk about 'past-dependent' innovation persistence. However, also 'path-dependent' processes are affected by context factors that influence the rate and direction of innovative processes in different periods and different locations.

exhaustible. This implies that firms that have been able to start creating new technological knowledge use their own knowledge stock to create new additional knowledge at a lower cost compared to competitors at the same time as they develop their innovative capability exploiting dynamic economies of scale.

Not only R&D but also learning is a major source of new knowledge (Arrow, 1962 a), which implies that the performance of innovative activities are influenced by dynamic increasing returns in the form of learning-by-doing effects, which increase knowledge stocks and the probability of successful future innovation. Previous innovation extends the firm's knowledge base and supplies knowledge inputs for future learning and knowledge creation, which may generate a virtuous cycle of innovation and knowledge creation. By innovating, a firm is engaged in a learning process through which it discovers new ideas be recombining existing ideas in new ways. The more knowledge pieces and ideas it has generated in the past, the higher is its ability to recombine them in order to generate new ideas and pieces of knowledge (Weitzman, 1996), which implies that past innovation affects current innovation (Duguet & Monjon, 2002). Thus, the larger the cumulative size of the innovation activities carried out, the lower are the innovation costs. An important dynamic element is the ability to learn to learn (Stiglitz, 1987), which implies that firms that have started to learn about how to create new knowledge can benefit from distinct dynamic increasing returns, since they are better able to learn in the subsequent attempts to generate new knowledge. For example, firms, which already have developed an experience and skills in efficiently cooperating with external knowledge and innovation partners, such as universities, consultancy firms and suppliers, are more likely to be successful in using external knowledge sources for future innovation projects than other firms. When knowledge, experience and learning ability accumulates over time the innovative performance of firms is augmented by idiosyncratic and non-imitable innovative competencies and capabilities (Kogut & Zander, 1992).

Frequent interactions between the accumulation of knowledge and the creation of routines to exploit that knowledge within the same organization may lead to the generation of dynamic technological capabilities that benefit the systematic reliance upon innovation as a competitive tool (Teece & Pisano, 1994). Such capabilities are a decisive factor in explaining innovation. The technological capabilities of firms are primarily determined by their level of human capital, i.e., by the knowledge, skills and creativity of their employees. Experience of innovation among the employees generate dynamic increasing returns as a result of learning effects, which increase a firm's knowledge stock and hence increase their technological and innovative capabilities. Furthermore, a firm's absorptive capacity is a function of the human capital of its employees and with increased learning in one period that further increases this absorptive capacity the firm will be able to more efficiently accumulate external knowledge in subsequent periods (Cohen & Levinthal, 1990). The cumulative nature of technological and innovative capabilities represents a process that might induce state dependence in innovation behaviour.

2.2 Sunk R&D Costs and Innovation Persistence

The knowledge creation process is also influenced by the sunk costs generated by earlier R&D investments (Sutton, 1991), which might induce state dependence, i.e. inter-temporal stability in firms' R&D efforts and innovation behaviour (Antonelli, Crespi & Scellato, 2012). The long-term commitments of firms to setup of R&D infrastructures and laboratories and the necessary long-term investments needed to be able to benefit from R&D returns are fixed outlays, which represent distinct sunk costs. The sunk cost hypothesis implies that firms deciding to invest in R&D incur start-up costs that usually are not recoverable except through the incomes from successful innovations. This implies that R&D investments over time generate a stock of physical and knowledge capital that in the longer term can be used in innovative activities and contribute to a more or less continuous flow of innovations. As R&D investments are a driver of innovation, the persistence of the former might lead to persistence of the latter, i.e. innovation (Cohen & Klepper, 1996).

Firms always face the choice between investing or not investing in R&D and innovation. However, decisions by firms to invest in R&D and to be active innovators necessitate the allocation of substantial resources for establishing, equipping and supporting R&D facilities, employment and training of specialized R&D staff and establishing advanced information systems for the collection and distribution of external and internal R&D results including patent applications as well as the implementation of the necessary routines (Máñez, et al., 2009). The effect is that as soon as the decision to innovate has been taken and the money has been spent the costs involved are sunk cost. This implies that the opportunity cost of ending the innovative activities are often quite high since the costs incurred mainly are unrecoverable, which indicates the high risks involved for firms engaging in innovative activities. A stop of the innovative activities also means that the dynamic increasing returns will be foregone. The combination of sunk costs and the irreversibility of R&D activities imply in R&D intensive industries that there are major entry and exit barriers. At the same time we have to observe that the presence of sunk costs reduce the costs of future innovative activities and thus induce innovating firms to continue innovating at the same time as it may prevent non-innovating firms to engage in innovative activities. Thus, it is natural that we in many industries should observe both innovating and non-innovating firms (Máñez, et al., 2009).

2.3 "Success breeds Success" and Resource Constraints

Successful innovative activities can be expected to have a positive impact on innovative firms' conditions for subsequent innovations by normally providing prosperous innovators with higher market power for an extended period, i.e. 'success breeds success' (Phillips, 1971). The innovation success of firms may broaden the space of available technological opportunities and opens up for exploiting economies of scope, which increases the probability of subsequent innovation success (Mansfield, 1968; Scellato & Ughetto, 2010).

Successful innovations also reduce the financial constraints of innovating firms partly because of increased market power. Resource constraints have been launched in the literature as an explanation of innovation persistence, which takes its starting point in the general observation that firms often meet serious financial limitations in financing their innovation projects. R&D and innovation ventures are often risky, capital-intensive and difficult for external financiers to assess (Arrow, 1962 b), which limits the possibility to use capital markets and other external sources of finance to get funding to finance innovation (Czarnitzki & Hottenrott, 2010) and instead force firms to finance them by means of internal funds. A stream of successful innovations provides firms with increased internal funding that can be used to finance innovations. It also lifts the external financing restrictions and makes banks and investors more interested and more willing to provide financing for ongoing innovative activities, since past success in innovation can be interpreted as an indicator of innovative capability and of possible future success in innovation. At the same time, it is natural that the incomes from earlier successful innovations can and will be used to finance current and future innovation. A bearing idea here is that firms launching commercially successful innovations gain a kind of lock-in advantage over less successful competitors.

2.4 Why Innovation May not be Persistent

The discussion above illustrates that there are many reasons why firms who have become innovators also will become persistent innovators. However, it is possible that counter-effects may exist, which induce some innovating firms to stop innovating. Thus, persistence would then not emerge or it would end. Firstly, we have the demand-pull case where a firm has perceptions of customers' demand that there is no need for further innovations due to their own previous innovations, which may induce a firm to stop innovating and instead concentrating on exploiting its earlier innovations (Schmookler, 1966). This might be the case for firms that offer only few products in markets characterized by rather long product cycles. Secondly, if an incumbent innovating firm fears that the introduction of further product innovations will cannibalise its rents from previous innovations it might be induced to stop innovating. Patent race models actually indicate that an incumbent firm who has innovated invests less in R&D than a challenging firm because further innovation might erode current monopoly profits (Reinganum, 1983). Thirdly, if the current product demand develops unfavourably, a firm might be forced for economic reasons to stop innovating, at least stop its product innovation.

2.5 What Induce Firms to Start Innovating?

One limitation of all the above explanations to innovation persistence fails to explain why firms at all start to invest in innovation. Why do firms start to invest in R&D with the hope of being able to introduce innovations in the market place? One obvious explanation could be that this would be one possible reaction when firms face unexpected events in factor or product markets. This implies that contextual factors matter to trigger off creative reactions within firms that may lead to the introduction of innovations.

When firms face such unexpected events, it is natural for them to try to mobilise and extend their internal knowledge stock through R&D and other learning processes. The probability that such a reaction will lead to an innovation is a function of the current internal knowledge stock, the ability to establish efficient R&D, learning and knowledge management routines and the capacity to search for and to absorb relevant external knowledge (Antonelli, 2011). This view implies that external knowledge networks, proximity to relevant knowledge sources and interaction with economic agents with varied knowledge bases are critical for a firm's ability to create new knowledge and to extend its knowledge stock. A strong reason why firms cluster is actually accessibility to a firm-relevant knowledge stock (Baptista & Swann, 1999). However, long-distance knowledge interactions can also be realised through organized proximity, not least within multinational firms (Torre & Rallet, 2005).

2.6 Persistence in Four Different Types of Innovations

Firms make innovations to improve their situation in the market place and their profit prospects but of course, their investments in different types of innovations are conditional upon the expected reactions of competitors. In line with Schumpeter (1934), we distinguish four main types innovation, namely, product, process, organizational and market innovation. A critical question here is if we shall expect equal persistence in all four types of innovation or not? Actually, the four types of innovation are not equal. However, all types of innovation demands organizational capabilities, even if the type of capabilities varies for the different types of innovation. Such capabilities are difficult to create and costly to adjust (Hannan & Freeman, 1984), which implies that when they have been created they tend to support persistence in innovation at the same time as they may make it difficult to shift between different types of innovation as well as between radical and incremental innovation. At the same time, firms have the strategic challenge to make product, process, organizational and market innovations work together to develop and preserve competitive advantage (Johne, 1999).

Naturally, we have strong reasons to expect that the existence of innovation persistence will vary between different industries and markets at a given point in time. Furthermore, starting from the assumption that there exist product cycles, we have reason to believe that the existence of persistence for the different types of innovation could vary over the life cycle of a product, firm and industry. Unfortunately, limitations in the available data imply that there are several hypotheses that we will not be able to test.

2.6.1 Product Innovations

Product innovations emerge when a new product or a new variety of an existing product is introduced in the market place for the first time aiming at satisfying a specific customer demand. Product innovations can but need not involve a technological innovation. This is obvious since products include both goods and services. A prime goal of product innovations is to introduce new products and new product varieties that allow the firm to gain at least a temporary monopoly position, which gives it a freedom to set prices above marginal costs. Given the critical role of product innovations for the long-term competitiveness of firms in many industries and markets, we assume that we will find the highest degree of innovation persistence for product innovations. Improved, radical changed and new products are conceived as particularly important for long-term firm growth and functions as a mean to help firms retain and grow their competitive position (Hart, 1996) and a fundamental condition for long-term market presence.

Firms that have adopted a product innovation strategy must to a high extent bind resources for a long time by setting up specialized R&D units in which the product innovation strategy is pursued. When such long-term investments have been made it is not meaningful to discontinue the R&D activities one year to take them up again the coming year, since that would be a waste of physical resources and of the R&D and innovation knowledge embodied in the researchers' human capital. This implies that when such investments have been taken, these firms are expected to have a continuous flow of product innovations. Certainly, firms increasingly out-source R&D work but that is probably in most cases a complement to the in-house R&D but can also be used as an alternative innovation strategy. If firms externally acquire the relevant knowledge and means for product innovation, we can expect that such a strategy lead to a less persistent innovation pattern (Verspagen & Clausen, 2012). In such a case, firms invest in one product innovation and then focus on exploiting this innovation in the market, without investing any further resources in product innovation.

2.6.2 Process Innovations

Process innovations involve the introduction of new methods of production, including new ways of handling a good or a service commercially. A primary goal for process innovations are the reduction of the unit costs of the products produced, which is achieved not least by introducing new machinery containing embodied knowledge. Other important goals are to preserve or increase the quality of the products produced. We must observe that, in particular, product innovations that involve the launching of completely new products may demand associated process innovations.

It is not clear-cut how one should distinguish process innovations from organizational innovations. However, we prefer to think that process innovations are associated with investments in new physical equipment embodying new knowledge, i.e. investments generating embodied technical change within the firm.

Concerning process innovation, we must acknowledge that such innovations differ from product innovations. In many industries most of the firms do not do major R&D to develop process innovations. Instead, machinery and process equipment is bought from firms in the machinery industries, who are specialised in developing and producing machinery and equipment that can be used for process innovations. In many industries, and in particular in process industries major process innovations are associated with the construction of totally new production units or factories such as paper machines and new pulp factories. Here process innovations involve large lumpy investments and we may not be able to observe persistence for (major) innovations. In these industries, it is not necessary to invest in large process R&D units, since the relevant research will be performed by the industries selling machinery and equipment for process industries.

2.6.3 Organizational and Market Innovations

Organizational innovations[5] are innovations involving changes in the routines of firms aiming at improving the efficiency, productivity, profitability, flexibility and creativity of a firm using disembodied knowledge. However, they often have the same goal as process innovations, namely to achieve cost reductions and quality improvements. Examples of such innovations are

1. Introduction and implementation of new strategies.
2. Introduction of knowledge management systems that improves the skills in searching, adopting, sharing, coding, storing and diffusing knowledge.
3. Introduction of new administrative and control systems and processes.
4. Introduction of new internal structures and types of work organization with their associated incentive structures including decentralized decision-making and team work.
5. Introduction of new types of external network relations with other firms and/or public organizations including, vertical cooperation with suppliers and/or customers, alliances, partnerships, sub-contracting, out-sourcing and off-shoring.
6. Mergers and acquisitions also fall within the category of organizational innovations.
7. Hiring of new personnel for key positions in the firm.

Market innovations involve the opening of new markets according to Schumpeter's classification but are in the modern management literature interpreted as improvements of the mix of target markets including market segmentation, and in methods to serve these markets (Johne, 1999). Innovations concerning the mix of markets include manipulation of the four famous marketing P's, i.e. product, price, promotion and place (including distribution methods and channels. This implies that the dividing line between product innovations and market innovations are not as clear-cut as one would wish. Primary goals here are to increase the total sales volume to make the exploitation of economies of scale possible to compete effectively with price, to effectively segment markets to catch a larger share of the consumer surplus and offer product characteristics and associated services that increase the willingness of customers to pay for these products. However, firms have to make a strategic choice between trying to supply (i) products at the lowest cost, (ii) products that are special in some way, or (iii) products

[5] Sometimes in the literature, organizational innovations are termed administrative innovations (Afuah, 1998).

focusing a distinct niche market, since firms cannot optimize their performance if they pursue different market strategies at the same time (Porter, 1985).

Organizational and market innovations are distinct from product innovations but have some resemblances with process innovations and with input innovations. Major organizational innovations are probably performed relatively seldom, since for organizations to function they need substantial periods to adapt to the new organization. The same is true for market innovations, since firms cannot confuse their customers with continuous changes marketing methods and promotions. Furthermore, when a firm has started to exploit a particular market, it often has limited resources to exploit simultaneously other markets. For organizational and market innovations firms and, in particular, small firms may rely on specialised consultancy firms to come up with the innovations, which implies that the firms don't always have to invest in large specialised units to carry through organizational and market innovations. Thus, we should expect a lower degree of innovation persistence in these two cases.

3. Empirical Evidence on persistency of innovation

Earlier empirical studies on innovation persistence used patent data as the measure of innovation and persistency of innovation. More recently, with the availability of Community Innovation Survey, it has become possible to measure innovation more directly and hence persistence studies used these data in various countries. Indeed it is argued that the panel data which is derived from innovation surveys reveals very different results to previous analyses of innovation persistence primarily based on patents data (Roper and Hewitt-Dundas, 2008; Peters, 2009). When it comes to estimation strategy, it seems the recently developed approach by Wooldridge (2005) become a method of choice in the empirical literature. We will use this approach and elaborate it in Section 6. The summary of major empirical studies dealing with persistency of innovation is presented in Table 1.

[Table 1 about here]

4. Data

The innovation related data in this study comes from five waves of the Swedish Community Innovation Survey (CIS) in 2004, 2006, 2008, 2010, and 2012. The CIS 2004 covers the period 2002-2004 and CIS 2006 covers the period 2004-2006 and so on, hence using the five ways, provide us with information about innovation activities of firms over a ten years period, i.e. from 2002 to 2012. In all five waves, there is information concerning product and process innovations as well as to innovation inputs (e.g. R&D investments). In the last three waves, there is also information concerning the marketing and organizational innovations. The survey consists of a representative sample of firms in industry and service sectors with 10 and more employees. Among them, the stratum with 10-249 employees has a stratified random sampling with optimal allocations and the stratum with 250 and more employees is fully covered. The response rates in the five waves vary between 63% and 86%, in which the later CIS waves having higher response rates compared with the earlier ones.

There are 21,105 observations in total, after appending all five waves of CIS[6]. Then we construct two panel datasets: (i) a balanced dataset consists of 2,870 observations, corresponding to 574 firms who participated in all five waves of CIS and (ii) an unbalanced dataset consists of 16,166 observations, corresponding to 4,958 firms participated in at least two consecutive waves (2,488 firms participated in two waves, 1,534 firms in three waves, and 936 firms in four waves). Finally, we merged the innovation-related data with other firm-characteristics data (e.g. export, import, ownership structure) coming from registered firm-level data maintained by Statistic Sweden (SCB). We use both panel and unbalance datasets in investigating state dependency (Section 5), while we only use panel dataset in investigating true state dependency, where we estimate a dynamic discrete choice model (Section 6). The definition of all variables is reported in the Appendix. The mean VIF score for all variables is 1.91 and each variable get a VIF score of below 3.5. This implies that multicollinearity is rather mild and may not bias the regression analyses results in the subsequent sections.

5. Is there a persistency in firms' innovation (state dependency)?

In order to investigate whether persistency exist or not (and if yes, to what extent), we used Transition Probabilities Matrix (TPM). TPM reveals the information about the probability of transitioning from one state to another. In our case, "state" is the innovation status of firms in each period, i.e. being an innovator (*INNO*) or being a non-innovator (*NON-INNO*). In par-

[6] This is obtained after the usual data cleaning, i.e. dropping observations with zero turnover or zero employees.

ticular, let a sequence of random variables $\{Y_1, Y_2, \dots, Y_n\}$ be a Markov chain. Then the TPM is formulated as follows:

$$\boldsymbol{TPM} = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1d} \\ p_{21} & p_{22} & \dots & p_{2d} \\ \vdots & \dots & \dots & \vdots \\ p_{d1} & p_{d2} & \cdots & p_{dd} \end{bmatrix} \tag{1}$$

Where,

$$p_{ij} = P(Y_t = j \,|Y_{t-1} = i) \tag{2}$$

Where p_{ij} measure the probability of moving from state *i* to state *j* in one period for the vector $\boldsymbol{Y}$. Finally, $\boldsymbol{Y}$ consists of several variables measuring different types of innovation, i.e. y_1 is product, y_2 is process, y_3 is marketing, and y_4 is organizational innovations. This TPM offers useful information for analysing persistence since it measures the probability that a firm goes from one state to another, while moving from one period to another period in time. p_{ij} are unknown parameters in our case and they can be estimated by Maximum Likelihood. It can be shown that the estimated parameters of p_{ij} equals to $\widehat{p_{ij}} = \frac{n_{ij}}{ni}$, where n_{ij} is the number of observed consecutive transitions from state *i* to state *j* and n_i is the total number of state *i*. In the context of innovation persistence, it is shown that persistency can exist in two forms of weak or strong (Cefis and Orsenigo, 2001; Roper and Hewitt-Dundas, 2008). First, there is a weak innovation persistency if sum of diagonal elements of the matrix ***TPM*** ($\mathrm{p_{ij}}, \mathrm{if\ i = j}$) is equal or bigger than 100% probability *but* not all elements of the diagonal of the matrix are equal to or higher than 50%. Second, there is a strong innovation persistency if sum of diagonal elements of the matrix ***TPM*** ($\mathrm{p_{ij}}, \mathrm{if\ i = j}$) is equal or bigger than 100% probability *and* all elements of the diagonal of the matrix ***TPM*** equal to or higher than 50%. Using ***TPM***, one can also calculate the Unconditional State Dependence (*USD*) as follows:

$$USD = p_{jj} - p_{ij} = P(Y_t = j \,|Y_{t-1} = j) - P(Y_t = j \,|Y_{t-1} = i) \tag{3}$$

Where, state *j* is *INNO* and state *i* is *NON-INNO*. *USD* is measured as Percentage Point (hereafter PP) and shows how much of the probability of being innovative in year *t* ($Y_t = j$) can be explained by the difference between being innovative ($Y_{t-1} = j$) versus being non-innovative ($Y_{t-1} = i$) in year *t-1*. *USD* is unconditional because it does not condition the state dependency on any observed or unobserved characteristics of the firm[7]. Table 2 reports the estimated parameters of Transition Probabilities Matrix as well as *USD*, using both balanced and unbalanced panel datasets.

[Table 2 about here]

Table 2 shows that there is a general pattern of strong persistency in innovative behaviour of firms, regardless of choosing balanced or unbalanced panel data sets. This is because the diagonal elements are usually above 50%. Since result of using balanced and unbalanced panels are similar, we will only discuss the result of balanced one for the sake of brevity. First, 77%[8] of innovative firms (could be any four types of innovation) persisted to stay innovative in the subsequent period, while only 23% shifted to become non-innovative. On the other hand, 60% of non-innovative firms also persisted to stay non-innovative in the subsequent period, while 40% shifted to become innovative. Moreover, the probability of being innovative in year *t+1* was about 37 PP higher for innovators than non-innovators in year *t* (37=77-40). This can be seen as a measure of unconditional state dependence[9]. Secondly, breaking down the innovative firms to the type of innovations they are engaging, Table 2 shows that there is also a general persistency pattern in all four types of innovations. However, as discussed in Section 2, the degrees of persistency in various types of innovation are not equal. In product innovation, 70% of the innovators in one year persisted in innovation in the subsequent year while 30% stopped their engagement. Moreover, the probability of being product innovator in year *t+1* was about 55 PP higher for product innovators than non-innovators in year *t*. In process innovation, 56% of the innovators in one year persisted in innovation in the subsequent year, while 44% stopped their engagement. Moreover, the probability of being product innovator in year *t+1* was about 31 PP higher for process innovators than non-innovators in year *t*. In marketing innovation, half of the innovators in one year persisted in innovation in the subsequent year, while the other half stopped their engagement. Moreover, the probability of being marketing innovator in year *t+1* was about 22 PP higher for marketing innovators than non-innovators in year *t*. Finally, in organizational innovation, 47% of the innovators in one year persisted in

[7] Other notations can be used for *USD*. For instance, Peters (2009) called it Observed State Dependence (OSD).
[8] This probability is obtained as follows: dividing 1093 transitions (that had innovation status as INNOVATIVE in year *t* and year *t+1*) by 1428 transitions (that had innovation status as INNOVATIVE in year *t*).
[9] This measure is an unconditional state dependence, since we have not controlled neither observed nor unobserved characteristics of firms yet. Therefore, we do not know yet how much of this state dependence is “true” or alternatively “spurious”. We will deal with it by incorporating the conditional state dependence in Section 6.

innovation in the subsequent year, while 53% stopped their engagement. Moreover, the probability of being organizational innovator in year *t+1* was about 24 PP higher for organizational innovators than non-innovators in year *t*. To sum up, among the various types of innovation, product innovators show relatively higher persistency in staying innovative in compare with other types of innovation (higher state dependence). Then process and marketing innovators are persistent in their innovative behaviour more or less with the same transition probabilities. Finally, organizational innovators are the least persistent innovators compared with other types of innovation. They could be seen as an exception the general pattern of strong persistency among various types of innovations. These firms indeed do not show strong persistency to staying organizationally innovative (47%). Nevertheless, they still show weak innovation persistency, since the sum of diagonal elements exceed 100% (77%+47%=124%). Such variation in the degree of persistencies in various types of innovation is what we expected and elaborated in Section 2.

6. Is there a true persistency in firms' innovation (true state dependency)?

6.1. Estimation Strategy

We employed a dynamic probit model in order to investigate the determinants of persistency of firms' innovation. Such model is able to analyse the conditional state dependence, hence allows us to distinguish between "true" state dependence from "spurious" one. This is necessary to do because the preliminary evidence of persistency found in Section 5 maybe (at least in part) due to observed and observed heterogeneity in firm's characteristics, i.e. spurious state dependency. The starting point is to assume that firm *i* invests in innovation activities in period *t* if the expected present value of profits happening to the investment in $y*_{it}$ is positive. The latent variable $y*_{it}$ depends on the previous and realized innovation $y_{i,t-1}$, observable vector of explanatory variables $\boldsymbol{X}_{it}$, and unobservable time-invariant firm-specific elements τ_i. Other time-varying unobservable elements are captured in the idiosyncratic error ε_{it}. Such relation can be formulated as follows:

$$y_{it}^{*} = \gamma y_{i,t-1} + \beta \boldsymbol{X}_{it} + \tau_i + \varepsilon_{it} \qquad (4)$$

If the latent $y*_{it}$ is positive then we observe that firm *i* introduces innovations, that is $y_{it} = 1$, and 0 otherwise. Furthermore, there are good reasons to believe that many firms in our sample do not start their innovation processes in the beginning of the period of this study, i.e. 2002. This means that the initial condition,y_{i0}, is presumably correlated with unobservable time-invariant firm-specific elements τ_i, leading to inconsistent estimators, known as initial condition problem. Moreover, it is possible that explanatory variables, X_{it}, are also correlated with τ_i (Ganter and Hecker, 2013; Antonelli et al, 2013). If these individual effects and the initial conditions are not properly accounted for, then the coefficient of the lagged dependent variable can be overestimated (Peters, 2009; Raymond et al, 2010). In order to accommodate such situation, Wooldridge modifies the original procedure of Heckman (1981a) by suggesting to model the distribution of $\{y_{i0}, \dots, y_{iT}\}$ given y_{i0} and to use Conditional Maximum Likelihood (CML) estimator (Wooldridge, 2005). Applying this approach, the time-invariant firm-specific elements can be decomposed as:

$$\tau_i = \alpha_0 + \alpha_1 y_{i0} + \alpha_2 \boldsymbol{X}_i + \alpha_i \tag{5}$$

Where $\boldsymbol{X}_i = \{\boldsymbol{X}_{i1}, \dots, \boldsymbol{X}_{iT}\}$ is the vector of explanatory variables in each period from t=1 to t=T and $\alpha_i \sim N(0, \sigma_a^2)$, which is assumed to be independent of y_{i0} and $\boldsymbol{X}_i$. Plugging (5) in (4), the probability that firm *i* introduce an innovation in period *t* can be formulated as follows:

$$Prob(y_{it} = 1 | y_{i0}, \dots, y_{i,t-1}, \boldsymbol{X}_{it}, \boldsymbol{X}_i, \alpha_i) = \boldsymbol{\phi}(\gamma y_{i,t-1} + \beta \boldsymbol{X}_{it} + \alpha_0 + \alpha_1 y_{i0} + \alpha_2 \boldsymbol{X}_i + \alpha_i) \tag{6}$$

Where y_{it} is a dichotomous variable getting value 1 if a firm *i* introduces innovation in year *t*. We operationalize introducing innovation in four ways: product, process, marketing, and organizational innovation. This way, we distinguish between four types of innovation rooted in Schumpeter's definition; hence, we have four different dependent variables. The parameter γ shows the effect of previous innovation on the probability of future innovation, i.e. persistency in innovation behaviour. $\boldsymbol{\phi}$ is the standard normal cumulative distribution function and $\boldsymbol{X}_{it}$ composed of observable firm characteristics: size, innovation input, physical capital, human capital, import, export, ownership structure, cooperation, and continuous R&D strategy (refer to Appendix for exact definition of each variable).

The main advantage of this estimator is that marginal effects can be estimated which is not possible in semi-parametric approaches. This allows us not only to determine whether true state dependence exists by referring to the significance level but also to highlight the magnitude of this phenomenon (if any) (Peters, 2009). τ_i is an unknown parameter, nevertheless, it can be estimated if we assume that it can gets its average value. Then, the Marginal Effects at Means (MEMs) of binary variable $y_{i,t-1}$ can be estimated as follows:

(7)

$$\widehat{MEMs} = \boldsymbol{\phi}\left(\hat{\gamma} + \widehat{\boldsymbol{\beta}}\boldsymbol{X^o} + \widehat{\alpha_0} + \widehat{\alpha_1}\bar{y}_{i0}\right) - \boldsymbol{\phi}(\widehat{\boldsymbol{\beta}}\boldsymbol{X^o} + \widehat{\alpha_0} + \widehat{\alpha_1}\bar{y}_{i0})$$

Where $\boldsymbol{X^o}$ is the vector of explanatory variables which is a fixed value that needs to be chosen (we used the mean values for all variables across *i* and *t*). Moreover, $\widehat{\boldsymbol{\beta}}$, $\widehat{\alpha_0}$, and $\widehat{\alpha_1}$ are the estimated parameters in Equation (6). The marginal effect estimated by Equation (7) shows the magnitude of the true state dependency or in other words, conditional state dependency.

6.2. Estimation Results

Table 3 reports the estimation results of random effect dynamic probit models in order to investigate the possible true state dependency in persistency of various types of innovations. The random effect probit model (elaborated in Section 6.1) assumes the strict exogeneity of explanatory variables. This is a strong assumption, because, for instance, it rules out the feedback effect between the future innovation introductions and size or R&D investment of firms. In order to assess the impact of including the explanatory variables, which may potentially fail the assumption of strict exogeneity, we follow the Peters's (2009) strategy of step-wise procedure. This means we start by specifying an extremely parsimonious model, in which only lagged innovation, initial condition and time and industry dummies are included, i.e. models (1), (3), (5), and (7). Then we add explanatory variables and inspect whether this affect the estimated state dependence effect, i.e. models (2), (4), (6), (8). The results of the estimations are presented in Table 3.

[Table 3 about here]

Concerning product innovation, it can be said that even after accounting for firms' unobserved heterogeneity (Model (1)) and observed heterogeneity (Model (1) and (2)); past innovation has a behavioural effect on future innovation. Particularly Model (2) controls for initial conditions, observed, and unobserved heterogeneity. This allows interpreting the significant effect of past innovation on future innovation as a "true" state dependency. The results concerning process innovation (in Model (3) and (4)) are similar to product innovation, in terms of significance of past innovation. Here again, it is possible to interpret the significant effect of past innovation on future innovation as a true state dependency. Past marketing innovation has the significant effect on the future innovation in Model (5). However, this significance is vanished in Model (6), where we control for observed heterogeneity and initial conditions.

This shows that marketing innovation does not have true state dependency on future behaviour and hence no casual inference can be drawn. The result for organizational innovation is somewhat similar to marketing innovation. Nevertheless, in Model (8), the past organizational innovation barely shows the significant effect on future behaviour.

In order to interpret the magnitude of the effect (true state dependency) properly, we have estimated the Marginal Effect at Means (*MEMs*) using Equation (7)[10]. Furthermore, we have distinguished the marginal effects based on size classes of firms. The result is reported in Figure 1.

[Figure 1 about here]

Looking at the general pattern, Figure 1 shows that the effect of previous innovation on future innovation (persistency) is the strongest among the product innovators. Then it comes to process and organizational innovators and finally the least persistency effect is identified for market innovators. Such general pattern is in place regardless of firms' size (i.e. in all size classes). To be more specific we look at each innovation type separately. First, being a product innovator increase the probability of introducing product innovation in the next period by 10.5 PP to 17.3 PP depending on the size classes, while the average is 15.3 PP (considering all size classes together). This means in average, introducing product innovation in current period increase the chance of introducing again a product innovation in the next period by 15.3 PP, controlling for observed and unobserved heterogeneity in firms' characteristics. This is indeed the magnitude of true state dependency (or conditional state dependency). Furthermore, it is interesting to compare the magnitude of such true state dependency with the Unconditional State Dependency (*USD*). The *USD* to introduce product innovation in $t + 1$ was 55 PP higher for product innovators than for non-innovators in period t (referring to Table 2). Controlling for unobserved and observed characteristics, this difference reduces to 15.3 PP. This implies that nearly one-third (15.3/55=0.28) of the initially observed product innovation persistency (identified by *USD*) can be attributed to "true state dependence", while the rest is due to observed and unobserved characteristics.

[10] Alternatively, estimating Average Marginal Effect (AME) reveals more or less the same magnitude effects, albeit slightly lower compared with MEMs for most of the innovation types.

Second, being a process innovator increase the probability of introducing product innovation in the next period by 8.7 PP to 12.9 PP depending on the size classes, while the average is 12 PP. This means in average, introducing process innovation in current period increase the chance of introducing again a process innovation in the next period by 12 PP, controlling for observed and unobserved heterogeneity in firms' characteristics. Furthermore, more than one third (12/31=0.38) of the initially observed process innovation persistency (identified by *USD*) can be attributed to "true state dependence", while the rest is due to observed and unobserved characteristics.

Third, being an organizational innovator increase the probability of introducing organizational innovation in the next period by 8.3 PP to 13.9 PP depending on the size classes, while the average is 12 PP (same as process innovation). This means in average, introducing organizational innovation in current period increase the chance of introducing the same type of innovation in the next period by 12 PP, controlling for observed and unobserved heterogeneity in firms' characteristics. Furthermore, half (12/24=0.5) of the initially observed organizational innovation persistency (identified by *USD*) can be attributed to "true state dependence". Another interesting point is that in terms of persistency, organizational and process innovations show very similar pattern. An exception can be found in larger firms, where the persistency in organizational innovations seems slightly to overtake the process innovation. This could be, for instance, due to higher persistency of strategic decisions taken by management in larger firms.

Lastly, being a market innovator increase the probability of introducing product innovation in the next period by 4.5 PP to 6.6 PP depending on the size classes, while the average is 6 PP. This is in line with the lack of significant persistency in market innovation (Table 3). This simply means market innovators are the least persistent innovators in compare with other types. This is what we expected (elaborated in Section 2), since firms do not want to confuse their customers by persistency changing the positioning, pricing strategy, and packaging features of their products in the market.

Apart from the lagged innovation, that shows the persistency, some observable firm characteristics turn out to affect the future innovation significantly. First, innovation input positively affects all type of innovation. This is not a surprise since this variable has some elements that can act as the input for technologically related innovations (e.g. product innovation) and nontechnologically-related innovation (e.g. marketing innovation). The elements for the former are, for instance, internal and external R&D investments and the elements for the latter is investment in activities dealing with market introduction of an innovation. Second, doing continuous R&D positively affects product innovation, while it negatively affects organizational innovation. The former can be explained by absorptive capacity concept (Cohen & Levinthal, 1990), while the latter shows the allocation of scarce resources and the choice that firms make in their innovation strategy. Third, the export intensity of firm shows the positive effect on product innovation, which is in line with trade version of endogenous growth models predict that export contributes to innovation and growth (Grossman and Helpman, 1991). Finally, human capital positively affects product and organizational innovation, while physical capital affects product and process innovation.

So far, we have investigated the persistency of various types of innovation independently. However, a closer look to our data told us that indeed 57% of innovators in our sample introduce more than one type of innovation at a given point in time. This necessitates a robustness check to account for possible interdependencies between firm's decisions to introduce various types of innovation simultaneously (and therefore avoid the potential bias resulting from modelling these decisions separately)[11]. In order to perform such robustness check, we employ multivariate random effect probit model, which is based on GHK simulation method for maximum likelihood estimation. This model allows for correlated random effects and error terms between various types of innovation. The result of such estimation shows that our main findings concerning persistency pattern in various types of innovation (Table 3 and Figure 1) still holds[12].

7. Conclusions

In this paper we investigated whether persistency exist in innovation of firms. Following Schumpeter, we distinguished between four types of innovation, while employing a long panel of Community Innovation Survey, which enabled us to trace the innovative behaviour of firms in Sweden over a ten years period. First, using Transition Probability Matrix, we

[11] If a high correlation in error terms of various innovation equations exists, it implies complementarities between various types of innovation through unobservable effects. Multivariate probit model makes a tetrachoric correlation conditional on covariates.

[12] The result of such robustness check is available upon request.

found the persistency behaviour in all types of innovation. However, the degree of persistency is not equal among various types of innovation, among which product innovators turns out to be the strongest persistent innovators. Second, using dynamic probit models, we investigate whether the persistency pattern that we found (state dependency) is a true state dependency or a spurious one. It turns out that product, process and organizational innovation have the true state dependency, while market innovation has the spurious one. This is because after controlling for observed and unobserved heterogeneity in firms' characteristics, the persistency effect still remained in all types of innovation except marketing innovation. When it comes to the magnitude of such true state dependency, once again, product innovators are ranked the highest. Being a product innovator increase the probability of introducing product innovation in the next period by 10.5 PP to 17.3 PP depending on the firm's size classes, while the average of 15.3 PP. Among the few existing studies, Ganter and Hecker (2013) found similar magnitude (17.7 PP) using German data. Furthermore, we detect that 57% of innovators in our sample introduce more than one type of innovation at a given point in time. We have controlled for this phenomenon in our analysis. However, we think such issue deserves further investigation. For instance, do firms have persistency in doing "combined" innovation strategy (e.g., whether firms persist to do both process and organizational innovation simultaneously)? Does engagement in any types of innovation lead to other types of innovation in future?

Table 1. Recent empirical studies concerning the persistence of innovation

Study	Sample and Time	Innovation Activities	Methodology	Measure of Persistency	Finding
Geroski et al. (1997)	British manufacturing firms, 1969–1988	Patents granted by the US PTO	Duration dependence, Weibull model	Length of innovation spell	Low persistence
Cefis & Orsenigo (2001)	French, German, Italian, Japanese, British., and American manufacturing firms, 1978-1993	Patent applications at EPO	Transition Probability Matrix used in first- and second order Markov chains	Probability of remaining in the same state of patenting	Bimodality, i.e. both great innovators and non-innovators have a high probability to remain in their state, while persistence is much lower in the intermediate classes.
Cefis (2003)	British manufacturing firms, 1978-1991	Patent applications at EPO	Transition Probability Matrix used in first- and second order Markov chains	Probability of remaining in the same state of patenting	Bimodality
Martinez-Ros & Labeaga (2009)	Spanish manufacturing firms, 1990-1999	Binary variables for product and process innovation obtained from ESEE survey	Dynamic random effects probit model and Wooldridge (2005) method	lagged (t-1) product and process innovations	(1) Persistence in innovation increases at least 15% the probability to develop more innovations (2) The introduction of the alternative innovation increases the probability to innovate in a range from 2 to 4% (complementarities)
Peters (2009)	German manufacturing and service firms, 1994-2002	A binary variable for innovation *input* (sum of investment in six innovation activities)	Dynamic random effects discrete choice model and Wooldridge (2005)'s method	lagged (t-1) binary measure of innovation input	High persistency (true state dependency)
Raymond et al (2010)	Dutch manufacturing firms, 1994-2002 (4 waves of CIS)	(1) A binary, indicating whether a firm is a technological product or process (TPP) innovator. (2) Share of innovative sales	Dynamic type 2 Tobit model with Wooldridge (2005) method (accounting for individual effects and handling the initial conditions problem)	(1) lagged (t-1) TPP innovator (2) lagged (t-1) share of innovative sales	(1) True persistence in the probability of innovating in the high-tech industries and spurious persistence in low-tech. (2) Past innovation output intensity affects current innovation output intensity in high-tech, while it has no such effect in low-tech.
Clausen et al (2011)	Norwegian firms in industrial sector, 1995-2004 (3 waves of CIS)	Binary variables for product and process innovation obtained from CIS and R&D survey	Dynamic random effects probit model with Wooldridge (2005) method (accounting for individual effects and handling the initial conditions problem)	lagged product and process innovations	Differences in innovation strategies across firms are an important determinant of the firms' probability to repeatedly innovate.
Ganter & Hecker (2013)	German firms, 2002-2008 (3 waves of CIS). Only balanced panel is used.	Binary variables for product, process, and organizational innovation	(1) Dynamic random effects probit model with Wooldridge (2005) method (2) Bivariate dynamic random effects probit model (to assess the potential interrelatedness between the adoption of organizational and technological innovation.)	lagged (t-2) product, process, and organizational innovations	(1) True persistence of product innovation (new to market) (2) No true persistence of product (new to firm), process, and organizational innovations
Haned et al (2014)	French manufacturing firms, 2002-2008 (3 waves of CIS). Only balanced panel is used.	Binary variables for product, process, and organizational innovation	Dynamic random effects probit model with Wooldridge (2005) method	lagged (t-2) product, process, both product t and process, and organizational innovations	A positive effect of organizational innovation on persistence in technological innovation (complementarities)

Table 2-Transition Probabilities

		Unbalanced Panel			Balanced Panel		
		Innovation status in t+1		USD	**Innovation status in t+1**		USD
Types of Innovation	**Innovation status in t**	NON-INNO	INNO		NON-INNO	INNO	
All types	NON-INNO	65%	35%	39 PP	60%	40%	37 PP
	INNO	26%	74%		23%	77%	
Product	NON-INNO	87%	13%	50 PP	85%	15%	55 PP
	INNO	37%	63%		30%	70%	
Process	NON-INNO	78%	22%	29 PP	75%	25%	31 PP
	INNO	49%	51%		44%	56%	
Market	NON-INNO	75%	25%	29 PP	72%	28%	22 PP
	INNO	46%	54%		50%	50%	
Organizational	NON-INNO	78%	22%	24 PP	77%	23%	24 PP
	INNO	54%	46%		53%	47%	

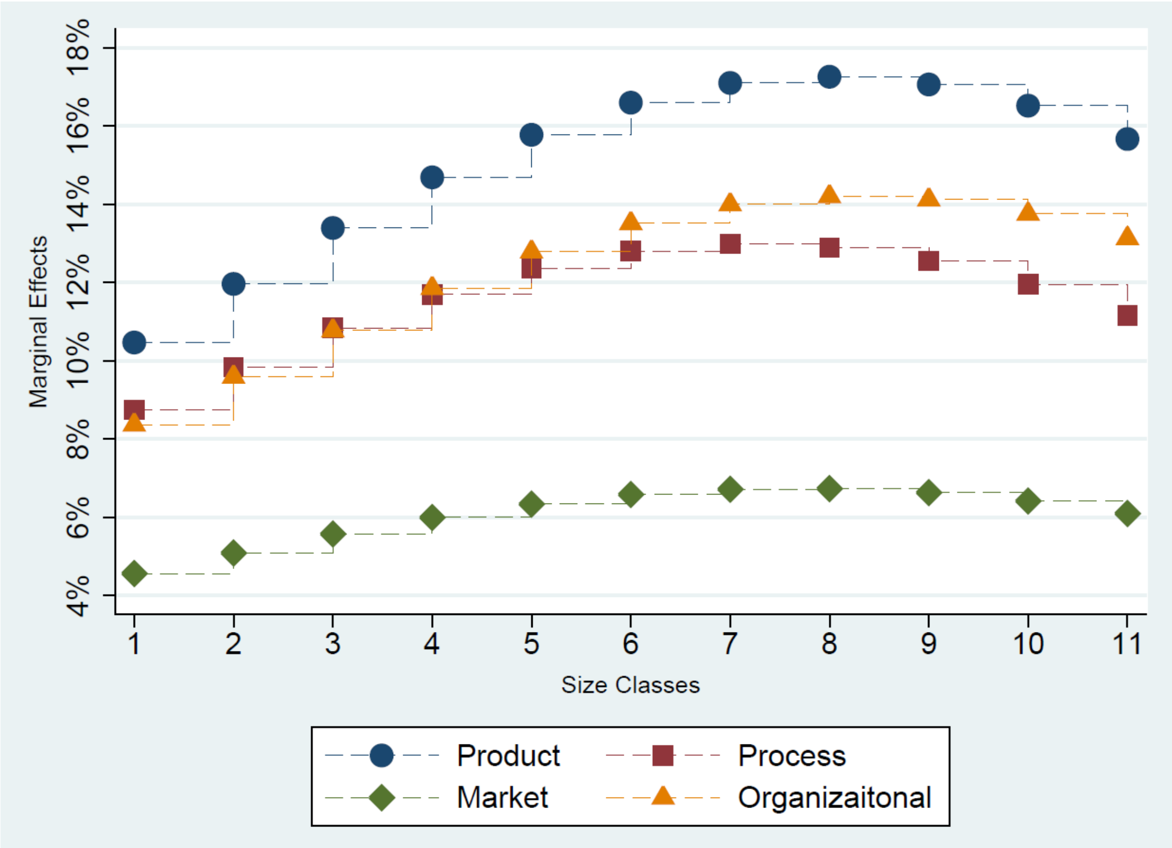
Notes: The table consists of ten 2x2 ***TPM*** matrices (five matrices under unbalanced panel and five under balanced panel). The table reports the estimated parameters of Transition Probabilities Matrices ($\widehat{p_{ij}} = \frac{n_{ij}}{ni}$). n_{ij} is the number of observed consecutive transitions from state i to state j and n_i is the total number of state i. Innovations status are the "state", which can be NON-INNO: Non-Innovative or INNO: Innovative. There are in total 10,644 transitions in the unbalanced panel and 2,296 transitions in the balanced panel. The sum of the rows in each matrix equals to 100%. The table also reports the *USD* (Unconditional State Dependence), as the Percentage Points (PP), which shows how much of the probability of being innovative in year *t* can be explained by the difference between being innovative versus being non-innovative in year *t-1*. t=2004, 2006, 2008, 2010, 2012.

Table 3- Dynamic Random Effect Probit models for various types of innovations

	$\mathrm{PRODUCT}_{it}$		$\mathrm{PROCESS}_{it}$		$\mathrm{MARKETING}_{it}$		$\mathrm{ORGANIZATIONAL}_{it}$	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\mathrm{PRODUCT}_{it-1}$	0.480***	0.354***						
	(0.115)	(0.127)						
$\mathrm{PRODUCT}_{i0}$	1.037***	0.688***						
	(0.145)	(0.131)						
$\mathrm{PROCESS}_{it-1}$			0.394***	0.199*				
			(0.089)	(0.102)				
$\mathrm{PROCESS}_{i0}$			0.503***	0.257***				
			(0.089)	(0.083)				
$\mathrm{MARKETING}_{it-1}$					0.353*	0.200		
					(0.191)	(0.188)		
$\mathrm{MARKETING}_{i0}$					0.218	0.142		
					(0.170)	(0.155)		
$\mathrm{ORGANIZATIONAL}_{it-1}$							0.456***	0.328*
							(0.177)	(0.179)
$\mathrm{ORGANIZATIONAL}_{i0}$							0.070	-0.067
							(0.158)	(0.143)
SIZE_{it-1}		0.062		0.049		0.080		0.117*
		(0.072)		(0.058)		(0.059)		(0.065)
$\mathrm{INNOV.\ INPUTS}_{it-1}$		0.016*		0.030**		0.027**		0.037***
		(0.015)		(0.013)		(0.013)		(0.013)
$\mathrm{COOPERATION}_{it-1}$		0.208		0.236		0.144		0.197
		(0.192)		(0.158)		(0.152)		(0.161)
$\mathrm{CONT.\ R\&}D_{it-1}$		0.399**		0.276		0.081		-0.363**
		(0.200)		(0.169)		(0.171)		(0.179)
IMPORT_{it-1}		-0.398		-0.463		-0.184		-0.055
		(0.515)		(0.461)		(0.402)		(0.421)
EXPORT_{it-1}		0.859***		0.062		0.211		0.037
		(0.302)		(0.244)		(0.242)		(0.254)
$\mathrm{PHYSICAL\ CAP}_{it-1}$		0.044*		0.067***		0.012		0.036
		(0.025)		(0.020)		(0.020)		(0.024)
$\mathrm{HUMAN\ CAP}_{it-1}$		1.059***		0.467		0.284		1.107***
		(0.386)		(0.291)		(0.362)		(0.384)
UNINATIONAL		-0.277*		-0.124		0.087		0.058
		(0.142)		(0.111)		(0.146)		(0.155)
DOMESTIC MNE		-0.206		-0.163		-0.195		-0.007
		(0.156)		(0.126)		(0.165)		(0.172)
FOREIGN MNE		-0.165		-0.304**		-0.313*		-0.206
		(0.164)		(0.131)		(0.176)		(0.184)
ρ	0.333	0.231	0.154	0.087	0.0051	0.008	0.085	0.005
	(0.059)	(0.066)	(0.048)	(0.051)	(0.148)	(0.139)	(0.138)	(0.143)
Log Likelihood	-1012.36	-945.67	-1330.94	-1257.69	-693.15	-663.80	-650.94	-609.83
Observations	2,296	2,296	2,296	2,296	1,140	1,140	1,140	1,140
Number of firms	574	574	574	574	574	574	574	574

Notes: The table reports the estimated parameters with standard errors in the parentheses. ***,** and * indicate significance on a 1%, 5% and 10% level. The estimation approach follows Wooldridge (2005). All models include sets of sector and time dummies. Models (2), (4), (6), (8) also include $\boldsymbol{x}_i$, which correspond to each of the explanatory variables in each period from t=2006 to t=2012. They are not shown in the table for the sake of brevity. Estimations are based on Gauss–Hermite quadrature approximations using twelve quadrature points. The accuracy of the results has been checked by applying eight, fourteen and sixteen quadrature points. ρ is the proportion of variance due to unobserved group level variance.

Figure 1-Marginal Effects for various types of innovations



Notes: The figure shows the marginal effects for four types of innovation over different size classes. Marginal effects are estimated as Marginal Effect at Means (MEMs) and shown in the above figure in terms of Percentage Points (PP). The estimation of MEMs for Product, Process, Marketing, and Organizational innovations are based on Models (2), (4), (6), (8) respectively, and employs Equation (7). Size classes is the logarithm of number of employments.

References

Aghion, P. & P. Howitt (1992), A Model of Growth through Creative Destruction, *Econometrica 60*, 323-351

Antonelli, C. (2011) (Ed.), *Handbook on Economic Complexity of Technological Change*, Edward Elgar, Cheltenham

Antonelli, C., F. Crespi & G. Scellato (2012), Inside Innovation Persistence: New Evidence from Italian Micro-Data, *Structural Change and Economic Dynamics 23*, 341-353

Antonelli, C., F. Crespi & G. Scellato (2013), Internal and External Factors in Innovation Persistence, *Economics of Innovation and New Technology 22*, 256-280

Arrow, K.J. (1962 a), The Economic Implications of Learning by Doing, *Review of Economic Studies 29*, 155-173

Arrow, K.J. (1962 b), Economic Welfare and the Allocation of Resources for Invention, in Nelson, R.R. (Ed.), *The Rate and Direction of Inventive Activity: Economic and Social Factors*, Princeton University Press, Princeton NJ, 609-625

Baptista, R. & G.M.P. Swann (1999), The Dynamics of Firm Growth and Entry in Industrial Clusters: A Comparison of the US and UK Computer Industries, *Journal of Evolutionary Economics 9*, 373-399

Cefis, E. (2003), Is There Persistence in Innovative Activities? *International Journal of Industrial Organization 21*, 489-515

Czarnitzki, D. & H. Hottenrott (2010), Financing Constraints for Industrial Innovation: What Do We Know?, *Review of Economics and Business 55*, 346-362

David, P.A. (1993), Knowledge Property and the Dynamics of Technological Change, in *Proceedings of the World Bank Annual Conference on Development Economics*, The World Bank, Washington DC

Fagerberg, J., D. Mowery & R.R. Nelson (2005) (Eds.), *The Oxford Handbook of Innovation*, Oxford University Press, Oxford

Flaig, G. & M. Stadler (1994), Success Breeds Success: The Dynamics of the Innovation Process, *Empirical Economics 19*, 55-68

Ganter, A. & A. Hecker (2013), Persistence of Innovation: Discriminating between Types of Innovation and Sources of State Dependence, *Research Policy 42*, 1431-1445

Geroski, P., J. Van Renen & C.F. Walters (1997), How Persistently Do Firms Innovate? *Research Policy 26*, 33-48

Grossman, G. M., & Helpman, E. (1991). Innovation and growth in the global economy. Cambridge: MIT Press.

Heckman, J. (1981 a), The Incidental Parameter Problem and the Problem of Initial Conditions in Estimating a Discrete Time-Discrete Data Stochastic Process, in Manski, C. & D. McFadden (Eds.), *Structural Analysis of Discrete Data*, The MIT Press, Cambridge MA, ?-?

Heckman, J. (1981 b), Heterogeneity and State Dependence, in Rosen, S. (Ed.),*Studies in Labor Markets*, University of Chicago Press, Chicago IL, ?-?

Kogut, B. & U. Zander (1992), Knowledge of the Firm, Combinative Capabilities, and the Replication of Technology, *Organization Science 3*, 383-397

Langlois, L.N. & N.J. Foss (1999), Capabilities and Governance: The Rebirth of Production in the Theory of Organization, *Kyklos 52*, 201-218

Máñez, J.A. et al. (2009), The Role of Sunk Costs in the Decision to Invest in R&D, *Journal of Industrial Economics 57*, 712-735

Nelson, R.R. (1959), The Simple Economics of Basic Scientific Research, *Journal of Political Economy 67*, 297-306

Penrose, E. (1959), *The Theory of the Growth of the Firm*, Oxford University Press, Oxford

Romer, P. (1990), Endogenous Technological Change, *Journal of Political Economy 98*, 71-102

Ruttan, V.W. (1997), Induced Innovation, Evolutionary Theory and Path Dependence: Sources of Technical Change, *The Economic Journal 107*, 1520-1529

Schumpeter, J. (1934), *The Theory of Economic Development*, The MIT Press, Cambridge, MA

Scellato, G. & E. Ughetto (2010), The Basel II Reform and Provision of Finance for R&D Activities, *International Small Business Journal 28*, 65-89

Stiglitz, J.E. (1987), Learning to Learn Localized Learning and Technological Progress, in Dasgupta, P. & P. Stoneman (Eds.), *Economic Policy and Technological Performance*, Cambridge University Press, Cambridge, ?-?

Sutton, J. (1991), *Sunk Costs and Market Structure*, The MIT Press, Cambridge MA

Teece, D. & C. Pisano (1994), The Dynamic Capabilities of Firms: An Introduction, *Industrial and Corporate Change 3*, 537-555

Torre, A. & A. Rallet (2005), Proximity and Localization, *Regional Studies 39*, 47-59

Wooldridge, Jeffrey M. (2005), Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity. *Journal of applied econometrics* 20.1: 39-54.

Appendix-Variable definitions

Variables	Type	Definitions
$PRODUCT_{it}$	0/1	1 if firm *i* introduces a product innovation into the market in year *t*, 0 otherwise. A product innovation is the market introduction of a new or significantly improved good or service with respect to its capabilities, user friendliness, components or sub-systems. Product innovations (new or improved) must be new to the enterprise, but they do not need to be new to the market.
$PROCESS_{it}$	0/1	1 if firm *i* introduces a process innovation in year *t*, 0 otherwise. A process innovation is the implementation of a new or significantly improved production process, distribution method, or support activity for goods or services, such as maintenance systems or operations for purchasing, accounting, or computing (exclude purely organizational innovation). Process innovations must be new to the enterprise, but they do not need to be new to your market.
$MARKET_{it}$	0/1	1 if firm *i* introduces a marketing innovation in year *t*, 0 otherwise. A marketing innovation is the implementation of a new marketing concept or strategy that differs significantly from the enterprise's existing marketing methods and which has not been used before. It requires significant changes in product design or packaging, product placement, product promotion or pricing. It exclude seasonal, regular and other routine changes in marketing methods.
$ORGANIZATIONAL_{it}$	0/1	1 if firm *i* introduces an organizational innovation in year *t*, 0 otherwise. An organizational innovation is a new organizational method in the enterprise's business practices (including knowledge management), workplace organization and decision making, or external relations that has not been previously used by the enterprise. It must be the result of strategic decisions taken by management. It exclude mergers or acquisitions, even if for the first time.
$INNOV\ INPUTS_{it}$	C*	Innovation inputs is the sum of following six expenditures in firm *i* year *t* (log): engagement in intramural R&D, engagement in extramural R&D, engagement in acquisition of machinery, engagement in other external knowledge, engagement in training of employees, and engagement in market introduction of innovation
$SIZE_{it}$	C	Number of employees in firm *i* year *t* (log)
$COOPERATION_{it}$	0/1	1 if firm *i* in year *t* had any cooperation with other customers, suppliers, competitors in, 0 otherwise
$CONT\ R\&D_{it}$	0/1	1 if firm *i* in year *t* had continuous R&D investments over the past two years, 0 otherwise
$IMPORT_{it}$	C	The amount (value in SEK) of import per employee for firm *i* in year *t* (log)
$EXPORT_{it}$	C	The amount (value in SEK) of export per employee for firm *i* in year *t* (log)
$UNINATIONAL_{i}$	0/1	1 if firm *i* belongs to a group and is uninational, 0 otherwise (Non-affiliated as based)
$DOMESTIC\ MNE_{i}$	0/1	1 if firm *i* belongs to group and is a domestic multinational enterprise, 0 otherwise
$FOREIGN\ MNE_{i}$	0/1	1 if firm belongs to group and is a foreign multinational enterprise, 0 otherwise
$PHYSICAL\ CAP_{it}$	C	Sum of investments in Buildings and Machines at year's end for firm *i* in year *t* (log)
$HUMAN\ CAP_{it}$	C	Share of employees with 3 or more years of university educations in firm *i* in year *t*
Time Dummies	0/1	Time-specific component captured by five time dummies
Sector Dummies	0/1	Sector-specific component captured by forty two sector dummies

*C corresponds to continuous variable